A study of intra-examiner error associated with recording of radiographic caries at different diagnostic levels

O. HAUGEJORDEN & G. L. SLACK

Dental Division, Directorate of Health, Oslo, Norway, and MRC Dental Epidemiology Unit, The London Hospital Medical College, England

Haugejorden, O. & Slack, G. L. A study of intra-examiner error associated with recording of radiographic caries at different diagnostic levels. *Acta Odont. Scand.* 33, 169—181, 1975.

Pairs of posterior bitewing radiographs of forty children 13—15 years of age at the third and fourth examinations of three-year clinical trials were assessed by one examiner for primary caries. Duplicate assessments were carried out in independent sessions, recording simultaneously absence/presence of caries and caries degree, and in the same way for three different diagnostic levels recording only absence/presence of caries.

Caries diagnostic inconsistency rate (D.I.R.) in per cent was always lower than 34 and usually below 20; a finding which is in agreement with previously published results. No preference could be expressed for any particular diagnostic level on the basis of the primary caries scores, but recording only relatively advanced lesions tended to give most favourable results for DFS increment scores.

The recording of both absence/presence of caries and caries degree simultaneously had a small, but not consistently adverse effect on the D.I.R. and reliability coefficient.

As there was no marked difference in examiner error associated with two of the demarcation lines used for recording of caries degree, i.e. half way through enamel and the A.D.J., the scoring codes described here may be suitable for studies of radiographic caries progression.

Key-words: Dental caries; diagnosis; radiography; dental

O. Haugejorden, Dental Division, Directorate of Health, Oslo-Dep., Oslo 1, Norway

Examples of descriptions of radiographic caries diagnostic criteria published in the literature were quoted by *Haugejorden* (1973). Some examiners recorded the earliest possible radiographic lesion (e.g. *Becks, Jensen & Millarr*, 1944; *James & Anderson*, 1967; *DePaola et al.*, 1968). Others required that the radiolucency should penetrate at least 1/3 through enamel (*Wittich*, 1950), or half way to the amelo-dentinal junction (A.D.J.) in posterior teeth (*Richardson, Hole & Williams*, 1971), while *Horowitz, et al.* (1966) only counted lesions which had reached the A.D.J. *Alman*'s (1965) consistency ratio (the formula is given later) for intra-examiner consistency of radiographic caries diagnosis using these different criteria varied between 70 and 90 per cent. It was, however, not possible on the basis

Received for publication, January 13, 1975.

Table 1. *Descriptions of scoring codes for recording of degree of radiographic approximal caries as reported in seven studies*

Authors + year of publication	Descriptions of numerical scoring codes '0'	'1'	'2'	'3'	'4'
Walker (1931)		when the enamel alone is involved	when the dentine is involved	when the pulp is involved	
Raper (1931)		those showing only a nick in the enamel	those showing a nick in the enamel and a radiolucent zone at the dentino-enamel junction	those showing a nick in the enamel with the zone in the dentine now bellying into the dentine towards the pulp	those showing a nick in the enamel, the perforation through the enamel and the carious process bellying well into the dentine
Emslie (1959)	caries free	caries of enamel only	with dentine involvement	gross dentine caries	
Backer Dirks (1966)	caries free	a radiolucent area that is limited to the enamel	the radiolucent area also visible in the dentine	the area extends beyond halfway between the dentino-enamel junction and the pulp chamber	the pulp chamber is reached
Hollender & Koch (1969)	caries free surface	carious lesion confined to the enamel	carious lesion involving both enamel and dentine	surfaces with restoration	
Berman & Slack (1973)	caries free	caries affecting enamel only and not reaching dentino-enamel junction	caries affecting enamel and just into dentine	caries affecting enamel and extending up to halfway through dentine	caries over halfway through dentine and to pulp
Erp & Meyer-Jansen *) (1970)	sound	initial lesion	dentine caries (cavity)	filling	extraction

*) A study of deciduous molars.

of these results to decide whether the recording of only relatively advanced lesions was associated with lower method error than when smaller lesions of less penetration were also counted, since inter-examiner variability among uncalibrated observers could have invalidated the conclusions. No conclusive evidence seems

to have been reported to show at what critical depth of penetration an approximal radiographic carious lesion should be recorded to be associated with minimum examiner error when radiographs are read for the absence/presence of caries.

Some investigators have not only recorded the absence/presence of caries but also the degree of caries at the same time. The thickness of enamel and dentine from the outline of the approximal surface to the pulp chamber as seen radiographically, has been subdivided to classify depths of penetration of the carious destruction (e.g. *Backer Dirks, van Amerongen & Winkler*, 1951; *Marthaler*, 1966; *Naylor & Emslie*, 1967; *Slack, et al.*, 1967a, 1967b). Examples of definitions of demarcation lines used to differentiate between various stages of development of radiographic carious lesions are summarized in Table I.

The extent of examiner error when using the demarcation lines between different degrees of caries as described in Table I, does not appear to have been reported. It would be useful to determine the magnitude of examiner error associated with recording of caries degree, as this would hopefully reveal which lines of demarcation might most advantageously be employed in studies of radiographic approximal caries progression.

The purpose of this investigation was, therefore, to determine the magnitude of intra-examiner diagnostic error,

i. when recording the absence/ presence of radiographic approximal caries at three diagnostic levels on a yes/no basis, and

ii. when using the same demarcation lines as above and recording absence/presence of caries, as well as, degree of caries at the same time.

MATERIAL AND METHODS

Forty pairs of right and left posterior bitewing radiographs taken at the third and fourth annual examinations in three-year clinical trials of caries preventive dentifrices were used. These radiographs had been exposed while using a modification of the film holder described by *Backer Dirks, et al.*, (1951) (*Slack et al.*, 1967a). Kodak DF 55 films had been used and exposure time and darkroom procedure had been according to the manufacturer's recommendations.

The two pairs of radiographs for each of the 40 subjects who were 13 and 15 years of age when these radiographs were taken, were mounted in transparent plastic film mounts which were assigned a code number for identification.

Aids to Radiographic Interpretation

A Watson's radiographic viewing box was used during interpretation of all radiographs. As only one of the working surfaces of the viewing box was in use, the other was covered to exclude extraneous light from the visual field. The light intensity on the viewing surface was constant, except for slight fluctuations which may have resulted from variations of the voltage of the mains electricity supply. A fixed magnifying lens ($\times$ 2) on the viewing box was used throughout.

A cardboard overlay was always used to ensure that the examiner assessed each pair of bitewing radiographs independently.

Method of Reading Radiographs

The longitudinal method of reading radiographs was used, i.e. the third and fourth sets of radiographs of each subject were read in chronological order in quick succession (*Haugejorden*, 1973).

Reading Conditions and Procedure

All interpretations were carried out by one examiner (O.H.) in a room with ordinary room lighting. The examiner's vision was checked and found to be within normal range.

Scores were assigned to approximal surfaces of premolars and permanent first and second molars. The reading procedure was as described by *Haugejorden* (1973) and the work was carried out at a comfortable unhurried pace.

Two carefully trained scribes were used in the study. They recorded the code number of each film mount on the scoring sheets before handing the mount to the examiner with the cardboard overlay in position. Simultaneous recording of the score by the scribe and by tape recorder permitted control of recording errors. All scoring sheets were checked against the tape recording, but only one or two mistakes were found.

The film mounts were arranged in two cardboard boxes containing 26 compartments each. Twenty-six of the forty film mounts were randomly selected and their arrangement in the first cardboard box determined by drawing lots. The remaining 14 sets were then randomly allocated position in the second cardboard box.

To allow checks of intra-examiner diagnostic consistency without the examiner being able to identify the sets of radiographs being reassessed, twelve sets were randomly selected from the 26 in the first cardboard box, and their position in the second box randomly determined. A break lasting 15 minutes after reading the film mounts in the first box, allowed the scribe time to transfer the 12 sets of radiographs which were due to be reassessed *in each recording* session from the first box to the empty compartments in the second box while the examiner was absent from the room. Thus the radiographs were assessed in the same order in every recording session.

About 22 months elapsed between recording of radiographs for absence/presence of caries and caries degree together, and solely for the absence/presence of caries on a yes/no basis. The interval of time between recording sessions was 1½—2 weeks, except when caries degree was also recorded, when it was about 20 days. The scores recorded at preceding recording sessions were never accessible to the examiner in subsequent sessions, and the time which elapsed between sessions made it impossible to remember scores previously assigned to particular surfaces.

Duplicate assessments in separate recording sessions were carried out for each diagnostic procedure to permit determination of intra- and inter-session intra-examiner error.

The scoring codes used in the first part of the investigation are described in Table II.

Score codes '1', '2', '3', and '4' were replaced by one code when caries degree was not recorded. The three different diagnostic levels employed were,

1) the smallest detectable radiographic carious lesion without minimum requirement of depth of penetration of the lesion,
2) the radiolucency had to penetrate at least half way into enamel for an approximal surface to be judged carious, and
3) the radiolucency had to be visible at the A.D.J. before 'caries' was recorded.

The order in which the three diagnostic levels were used, was randomly determined.

Table II. *Scoring codes for radiographic diagnosis*

Score code	Description
0 —	caries free surfaces
1 —	caries in the enamel not penetrating more than half way into enamel
2 —	caries penetrating more than half way into enamel but not involving the amelo-dentinal junction
3 —	caries of enamel and dentine extending less than half way into dentine
4 —	caries of enamel and dentine penetrating more than half way into dentine or apparently to the pulp cavity
5 —	filled surfaces without evidence of secondary caries
6 —	extracted surfaces (regardless of reason for extraction)
7 —	surfaces missing from the film or overlapping and unreadable
8 —	unerupted surfaces

Degree of secondary caries was recorded by adding a suffix to the primary caries scores, viz., 1_x, 2_x, 3_x, 4_x.

If in doubt regarding the existence of a carious lesion according to these criteria, then 'sound' was recorded. When caries degree was recorded, the lesser degree of caries was recorded when doubt arose about depth of lesion relative to the demarcation lines described in Table II.

The definitions of primary and secondary radiographic caries, state of eruption and readability status of approximal surfaces were those described by *Haugejorden* (1973). The only difference was that the recording of approximal surfaces as overlapped/unreadable was adjusted depending on the minimum depth requirement for recording of 'caries'. Thus when the radiolucency had to involve the A.D.J. before 'caries' was recorded, an overlap involving only enamel, but not the A.D.J. would be regarded as 'sound'.

Analysis of the Data

If c represents a carious surface, cc a surface judged to be carious at two independent assessments and s represents a sound or unreadable surface, then an examiner's diagnostic consistency can be calculated in per cent from the formula for the consistency ratio (C.R.) (*Alman*, 1965),

$$\text{C.R.} = \frac{\text{cc}}{\text{cc} + \text{sc} + \text{cs}} \times 100.$$

Haugejorden (1973) argued that it might be an advantage to focus attention on diagnostic inconsistency rather than consistency and gave examples to support the idea. He proposed rewriting Alman's formula in accordance with the suggestion. The rewritten formula gives the diagnostic inconsistency rate (D.I.R.) in per cent, i.e.

$$\text{D.I.R.} = 100 - \frac{\text{cc}}{\text{cc} + \text{sc} + \text{cs}} \times 100.$$

The formula shows that D.I.R. is zero per cent when all diagnostic decisions at two independent assessments are concordant (cc), and 100 per cent when all decisions are discordant (sc or cs).

For example, in Table III (first line) cc = 22, sc + cs = 1, therefore

$$\text{D.I.R.} = 100 - \frac{22}{22 + 1} \times 100 = 100 - 95.7 = 4.3,$$

i.e. 4.3 per cent.

A serious disadvantage of this and the conventional methods of expressing examiner inconsistency or consistency is that estimates cannot be used in the analysis of the data. If, however, as in this investigation, duplicate assessments of the radiographs are carried out, then it

Table III. *Number of surfaces according to recorded status at duplicate assessments, and the intra-session D.I.R.*

(N = 12 subjects)

Examination	Concordant	Discordant	Total	D.I.R. %
A. Recording of earliest detectable primary radiographic carious lesion				
First session				
Third	22	1	23	4.3
Fourth	29	2	31	6.5
Second session				
Third	22	3	25	12.0
Fourth	29	2	31	6.5
B. Recording of primary radiographic caries when penetrating at least half way through enamel				
First session				
Third	14	7	21	33.3
Fourth	25	3	28	10.7
Second session				
Third	18	4	22	18.2
Fourth	28	1	29	3.4
C. Recording of primary radiographic caries at the A.D.J.				
First session				
Third	9	3	12	25.0
Fourth	14	2	16	12.5
Second session				
Third	11	2	13	15.4
Fourth	13	2	15	13.3

becomes possible to calculate the measurement error(s_e),

$$s_e = \sqrt{\frac{\Sigma d^2}{2N}},$$

where d is the difference between duplicate scores for each set of radiographs and N is the number of differences (*Grainger*, 1967).

Since all sets of radiographs were assessed twice by each method, two estimates of total variance of the observed scores of the group can be obtained. Total variance minus error variance(s_e^2) gives an estimate of true variance. The mean of the two estimates of total variance is used here as total variance(s_t^2) and the estimated true variance(s_x^2) is taken to be $s_t^2 - s_e^2$. Every reasonable effort should be made to reduce error variance, i.e. to make the ratio between true and total variance as nearly equal one as possible. This ratio has been termed the reliability coefficient (r_{tt}) for a diagnostic method (*Guilford & Fruchter*, 1973),

$$r_{tt} = \frac{s_x^2}{s_t^2}.$$

The results of the investigation will be expressed in terms of the D.I.R., the reliability coefficient, and the measurement error as a percentage of total caries or DFS (D = decayed, F = filled, S = surface) increment score.

RESULTS

The intra-examiner D.I.R. in each recording session at the three different diagnostic levels are given in Table III. The D.I.R. varied between 4 and 12 per cent when the earliest detectable carious lesions were recorded on a dichotomous scale. When penetration of the lesion at least half way through enamel was required before inclusion of a lesion in the count, the D.I.R. was 3—33 per cent. Recording only lesions which had at least reached the A.D.J. gave D.I.R. values between 12 and 25 per cent. The corresponding D.I.R. values when caries degree was also recorded, varied from 13—18 per cent, 14—21 per cent, and 8—18 per cent, respectively (Table IV). The D.I.R. tended

Table IV. *Number of surfaces according to recorded status at duplicate assessments, and the intra-session D.I.R. when caries degree is recorded*

(*N = 12 subjects*)

Examination	Concordant	Discordant	Total	D.I.R. %
A. Recording of earliest detectable primary radiographic carious lesion				
First session				
Third	42	9	51	17.7
Fourth	54	8	62	12.9
Second session				
Third	37	7	44	15.9
Fourth	48	7	55	12.7
B. Recording inconsistency at the half way through enamel demarcation line				
First session				
Third	16	4	20	20.0
Fourth	27	7	34	20.6
Second session				
Third	16	3	19	15.8
Fourth	24	4	28	14.3
C. Recording inconsistency at the A.D.J.				
First session				
Third	9	1	10	10.0
Fourth	12	1	13	7.7
Second session				
Third	9	2	11	18.2
Fourth	10	2	12	16.7

Table V. *Number of surfaces according to recorded status at duplicate assessments, and the inter-session D.I.R.*

(*N = 40 subjects*)

Examination	Concordant	Discordant	Total	D.I.R. %
A. Recording of earliest detectable primary radiographic carious lesion				
Third	60	7	67	10.4
Fourth	75	11	86	12.8
B. Recording of primary radiographic caries when penetrating at least half way through enamel				
Third	44	12	56	21.4
Fourth	64	7	71	9.9
C. Recording of primary radiographic caries at the A.D.J.				
Third	25	4	29	13.8
Fourth	31	6	37	16.2

to be slightly higher when assessing third than fourth examination radiographs, except in the first recording session when recording the earliest detectable lesions, where the reverse applied.

Inter-session D.I.R. values were of the same order of magnitude as the intra-session figures quoted earlier (Tables V and VI). The range of observed D.I.R. percentages was narrower between sessions than within sessions. The D.I.R. values had also a tendency to be slightly higher in Table VI than in Table V, i.e. when absence/presence of caries and caries degree were recorded simultaneously.

When the second approach to analysis was used, the results for the three different diagnostic levels were as shown in Tables VII and VIII. The reliability coefficient for primary caries scores varied between 0.933 and 0.988 when recording only absence/presence of radiographic caries, and 0.894 and 0.966 when caries degree was also recorded. The measurement error expressed in per cent of total primary caries score ranged from 0.39—1.09 and 0.49—1.35, respectively (Tables IX and X). Neither the reliability coefficient nor the measurement error expressed in per

Table VI. *Number of surfaces according to recorded status at duplicate assessments, and the inter-session D.I.R. when caries degree is recorded*

(N = 40 subjects)

Examination	Concordant	Discordant	Total	D.I.R. %
A. Recording of earliest detectable primary radiographic carious lesion				
Third	106	20	126	15.9
Fourth	128	26	154	16.9
B. Recording inconsistency at the half way through enamel demarcation line				
Third	42	11	53	20.8
Fourth	56	17	73	23.3
C. Recording inconsistency at the A.D.J.				
Third	20	7	27	25.9
Fourth	28	6	34	17.6

Table VII. *Total, true and error variance, and reliability coefficient (r_{tt}) for primary radiographic caries scores according to examination*

(N = 40 subjects)

Characteristic	Caries diagnostic level 'Earliest'	'Half way'	'A.D.J.'
Third examination			
Total var.	5.957	4.113	1.379
True var.	5.727	3.838	1.329
Error var.	0.230	0.275	0.050
r_{tt}	0.961	0.933	0.964
Fourth examination			
Total var.	7.934	6.163	2.233
True var.	7.504	6.088	2.183
Error var.	0.430	0.075	0.050
r_{tt}	0.946	0.988	0.978

Table VIII. *Total, true and error variance and reliability coefficient (r_{tt}) for primary radiographic caries scores when caries degree is recorded*

(N = 40 subjects)

Characteristic	Caries diagnostic level 'Earliest'	'Half way'	'A.D.J.'
Third examination			
Total var.	11.038	2.863	0.965
True var.	10.663	2.688	0.877
Error var.	0.375	0.175	0.088
r_{tt}	0.966	0.939	0.909
Fourth examination			
Total var.	13.510	3.441	1.827
True var.	13.022	3.078	1.764
Error var.	0.488	0.363	0.063
r_{tt}	0.964	0.894	0.966

Table IX. *Measurement error in per cent of total primary caries score according to diagnostic level and recording session when only absence/presence of caries was recorded*

(N = 40 subjects)

Caries diagnostic level	Measurement error (s_e)	Recording session First Total score	First s_e %	Second Total score	Second s_e %
Third examination					
'Earliest'	0.48	62	0.77	63	0.76
'Half way'	0.52	48	1.09	52	1.01
'A.D.J.'	0.22	27	0.83	27	0.83
Fourth examination					
'Earliest'	0.66	81	0.81	78	0.85
'Half way'	0.27	67	0.40	69	0.39
'A.D.J.'	0.22	34	0.65	34	0.65

Table X. *Measurement error in per cent of total primary caries score according to demarcation line and recording session when caries degree is also recorded*

(N = 40 subjects)

Caries diagnostic level	Measurement error (s_e)	Recording session: First Total score	First s_e %	Second Total score	Second s_e %
		Third examination			
'Earliest'	0.61	122	0.50	110	0.56
'Half way'	0.42	49	0.85	45	0.93
'A.D.J.'	0.30	22	1.35	25	1.18
		Fourth examination			
'Earliest'	0.70	143	0.49	140	0.49
'Half way'	0.60	68	0.89	59	1.02
'A.D.J.'	0.25	31	0.81	30	0.83

Table XII. *Measurement error in per cent of total DFS increment score according to diagnostic level/demarcation line and recording session*

(N = 40 subjects)

Caries diagnostic level	Measurement error (s_e)	Recording session: First Total score	First s_e %	Second Total score	Second s_e %
		Recording only absence/presence of caries			
'Earliest'	0.66	26	2.54	23	2.87
'Half way'	0.65	26	2.50	26	2.50
'A.D.J.'	0.34	16	2.12	15	2.27
		Recording both absence/presence and caries degree			
'Earliest'	0.81	30	2.71	37	2.20
'Half way'	0.71	28	2.54	24	2.96
'A.D.J.'	0.35	19	1.84	13	2.69

Table XI. *Total, true and error variance, and reliability coefficient (r_{tt}) for DFS increment scores according to diagnostic level*

(N = 40 subjects)

Characteristic	Caries diagnostic level: 'Earliest'	'Half way'	'A.D.J.'
Recording only absence/presence of caries			
r_{tt}	2.242	3.515	0.987
Total var.	1.812	3.090	0.874
True var.	0.430	0.425	0.113
Error var.	0.808	0.879	0.886
Recording both absence/presence and caries degree			
Total var.	4.361	2.718	0.702
True var.	3.698	2.218	0.577
Error var.	0.663	0.500	0.125
r_{tt}	0.848	0.816	0.822

cent of total caries score exhibited a consistent tendency towards more favourable results when only the more advanced lesions were counted.

As the DFS increments are commonly used to assess the effectiveness of potentially caries prophylactic agents or procedures, it was considered useful also to determine the reliability coefficient and the measurement error for DFS increment scores.

The reliability coefficient for DFS increment scores were 0.808—0.886 when recording absence/presence of caries, and 0.816—0.848 when the degree of caries was also recorded (Table XI). The measurement error in per cent of total DFS increment score varied from 2.12—2.87 when recording absence/presence of caries, and from 1.84—2.96 when recording caries degree as well (Table XII). There was a tendency for the percentages in Table XII to be higher when referring to recording of the earliest detectable

radiographic lesion and lowest when the diagnostic level was set at the A.D.J. The values obtained when stipulating that a lesion must penetrate at least half way through enamel before it is counted, were intermediate. This trend was, however, not consistent for the reliability coefficient (Table XI).

DISCUSSION

All estimates, except one, i.e. D.I.R. equals 33.3 per cent, were within or below the range of 20—27 per cent calculated from other published results concerning intra-examiner error when recording radiographic caries (*Hollender & Koch*, 1969, *Berman*, 1970, *Richardson, et al.*, 1971, *DePaola & Alman*, 1972; *Haugejorden*, 1973).

The lower intra-session D.I.R. values resulting when scoring the earliest detectable radiographic carious lesion (Tables V and VI) could be due to the shorter list of scoring codes which omission of registration of caries degree entailed. A more likely explanation for the observed difference is the change to a lower recording level which occurred in the 22 months between recording for absence/presence of caries plus caries degree, and recording only absence/presence of caries. This was the case especially when recording the earliest detectable lesion. Another possible reason for the differences in intra-session caries diagnostic inconsistency is the small number of diagnostic decisions on which the percentage D.I.R. is based. Finally, variation in the examiner's day-form cannot be ruled out as a contributing factor even though the examiner took care to be well rested for each session and took every reasonable precaution to avoid distractions or interruptions while recording.

There are four possible explanations for the tendency for inter-session D.I.R. values to be somewhat higher when caries degree was also recorded than when it was not (Tables V and VI). Firstly, it could be due to variation in the examiner's day-form; secondly, fluctuations in the examiner's caries recording level on the same criteria from one session to the next; thirdly, the use of a more complicated set of scoring codes may have played a part; and fourthly, the interval of time between recording sessions was seven days longer when caries degree was also recorded. The fourth reason for the observed difference may, however, have played a minor part, as *Haugejorden* (1973) found that *Alman*'s (1965) consistency ratio did not tend to decrease as the period of time between sessions increased.

The reliability coefficient, and measurement error expressed in per cent of total primary caries and DFS increment scores showed the same trends as the D.I.R., but no data was discovered which allowed direct comparison with the results presented in Tables VII—XII. *Rugg-Gunn* (1972) has, however, reported reliability coefficients for DFS scores recorded by clinical examination at two diagnostic levels. His results were similar to those reported for primary radiographic caries, but as diagnosis of the F-component of the DFS index is associated with minimal error, there may be greater reliability in radiographic than in clinical caries diagnosis. However, this suggestion needs to be subjected to experimental testing since two different examiners used the clinical and radiographic examination techniques.

The reliability coefficient was lower for DFS increments than for primary caries for both recording procedures and for all three diagnostic levels, i.e. 0.81 to

0.89 and 0.89 to 0.99, respectively. The most likely reasons for this consistent trend is, (1) examiner error, (2) lack of reproducibility of the radiographic technique (*Kyes, Overton & McKean*, 1961; *Haugejorden*, 1973); (3) remineralization of subsurface carious lesions (*Torell*, 1967) making them radiographically invisible or less well defined and hence more difficult to diagnose consistently; and (4) changes caused by tooth eruption, extractions, migration or rotation and by insertion of restorations. How much of the measurement error for DFS increment scores was due to each of the factors just mentioned, could not be determined from these results. However, some indication of the importance of the radiographic technique was given by *Haugejorden* (1973) who found that about 50 per cent of the reversal rate was due to score transitions from carious to unreadable. It follows from these observations that measurement error for DFS increment scores can be expected to be lower when studying groups with stable dentitions and no access to dental treatment. It follows also that measurement error for DFS increment scores should tend to be lower for control than for test groups in studies of caries preventive agents which facilitate remineralization of subsurface enamel lesions. Findings in support of this contention have been reported (*Muhler*, 1959; *Radike*, 1960).

Table I shows that the half way through enamel demarcation line for distinguishing degrees of caries has not been employed before in studies of caries progression. The results presented suggest that it may be possible to use this, and the A.D.J. demarcation line as defined here when studying approximal radiographic caries progression, since the intra-examiner caries diagnostic inconsistency at the half way through enamel demarcation line was not markedly different from that observed for the A.D.J. or when recording the earliest detectable radiographic lesion.

Six of the eight estimates of measurement error in per cent of total primary caries score (Tables IX and X) support *Backer Dirks'* (1964) statement that the relative size of the method error is smaller if the caries diagnosis is made at an earlier rather than a later stage of the disease process. One of the four corresponding estimates for DFS increment scores showed the same tendency (Table XII). Thus the results do not identify one stage of development of the carious lesion at which the measurement error is consistently lowest, and the ability of the examiner to diagnose caries consistently will be an important consideration when deciding on the appropriate recording level.

Though the results tend to show that examiner diagnostic inconsistency was higher when absence/presence of caries and caries degree were recorded simultaneously than if recording was limited to absence/presence of caries, the fact that some differences were in the opposite direction to expectations, means that no firm conclusion can be drawn on the basis of these results. The differences observed for this examiner were, however, so small that the savings in labour and cost made possible by recording of absence/presence and caries degree simultaneously would in most cases be more important than the difference in examiner error.

CONCLUSIONS

The results justify the following conclusions:

1) This examiner's diagnostic error was not improved by limiting the scoring to only relatively advanced primary carious lesions.

2) Diagnostic error was slightly, but not invariably adversely affected by simultaneous recording of absence/presence of caries and caries degree compared with the results obtained when omitting registration of caries degree. The differences were too small from a practical point of view to justify separate recording of absence/presence of caries and caries degree in studies in which both caries incidence and progression are of interest.

3) When training examiners for scoring of primary radiographic caries, it seems reasonable to aim at attaining D.I.R. values lower than 20 per cent or reliability coefficients higher than 0.90.

4) The measurement error for recording DFS increments was in absolute terms about halved when recording only lesions visible at the A.D.J. as compared with the values found when registering the earliest detectable lesions, but relative to total DFS increment score for the group the advantage was slight. This finding is probably primarily due to the almost total absence of score transitions from carious to unreadable when recording only lesions visible at the A.D.J.

5) Further research is indicated to determine the magnitude of the contribution which factors other than examiner inconsistency make to total measurement error for DFS increment scores.

REFERENCES

Alman, J. E. 1965. Notes on models for evaluation of examiner error in dental studies. Document 8/31/65, p. 1—8

Backer Dirks, O. 1964. At what stage should a carious lesion be recorded. In: *Advances in fluorine research and dental caries prevention.* Vol. 2, Oxford, Pergamon Press, pp 21—23

Backer Dirks, O. 1966. Posteruptive changes in dental enamel. J. Dent. Res. 45, 503—511

Backer Dirks, O., van Amerongen, A. J. & Winkler, K. C. 1951. A reproducible method for caries evaluation. J. Dent. Res. 30, 346—359

Becks, H., Jensen, A. L. & Millarr, C. B. 1944. Rampant dental caries: Prevention and prognosis. A five year clinical survey. J. Amer. Dent. Ass. 31, 1189—1200

Berman, D. S. 1970. A study of the pattern of development of dental caries of the permanent dentition in selected groups of children. Ph.D. Thesis, University of London

Berman, D. S. & Slack, G. L. 1973. Caries progression and activity in approximal tooth surfaces. A longitudinal study. Brit. Dent. J. 134, 51—57

DePaola, P. F. & Alman, J. 1972. Assessment of the reliability of radiographic diagnosis in a clinical caries trial. J. Dent. Res. 51, 1431—1437

DePaola, P. F., Wellock, W. D., Maitland, A. & Brudevold, F. 1968. The relationship of cariostasis, oral hygiene, and past caries experience in children receiving three sprays annually with acidulated phosphate-fluoride: three-year results. J. Amer. Dent. Ass. 77, 91—94

Emslie, R. D. 1959. Radiographic assessment of approximal caries. J. Dent. Res. 38, 1225—1226

Erp, N. A. K. M. van & Meyer-Jansen, A. C. 1970. A caries study of the temporary molars and its significance for their regular conservative care. Netherland Dent. J. 77, 51—74 (suppl. 5)

Grainger, R. M. 1967. Epidemiological data. In: *Design and analysis in dental and oral research.* By N. W. Chilton. Philadelphia, J. B. Lippincott Co. p. 317

Guilford, J. P. & Fruchter, B. 1973. *Fundamental statistics in psychology and education.* 5th ed. New York, McGraw-Hill, p. 399

Haugejorden, O. 1973. A study of the methods of radiographic diagnosis of dental caries in epidemiological investigations. Ph.D. Thesis, University of London, 269 p.

Hollender, L. & Koch, G. 1969. Influence of topical application of fluoride on rate of progression of carious lesions in children. Odont. Revy 20, 37—41

Horowitz, H. S., Law, F. E., Thompson, Mary B. Chamberlin, Shirley R. 1966. Evaluation of stannous fluoride dentifrice for use in dental public health programmes. I. Basic findings. J. Amer. Dent. Ass. 72, 408—422

James, P. M. C. & Anderson, R. J. 1967. Clinical testing of a stannous fluoride calcium pyrophosphate dentifrice in Buckinghamshire school children. Brit. Dent. J. 123, 33—39

Kyes, F. M., Overton, N. J. & McKean, T. W. 1961. Clinical trials of caries inhibitory dentifrices. J. Amer. Dent. Ass. 63, 189—193

Marthaler, T. M. 1966. A standardized system of recording dental conditions. Helv. Odont. Acta 10, 1—18

Muhler, J. C. 1959. The combined anticariogenic effect of a single stannous fluoride solution and unsupervised use of a stannous fluoride containing dentifrice. J. Dent. Res. 38, 994—997

Naylor, M. N. & Emslie, R. D. 1967. Clinical testing of stannous fluoride and sodium monofluorophosphate dentifrices in London school children. Brit. Dent. J. 123, 17—23

Radike, A. W. 1960. A study of reversals in diagnosis of carious lesions. In: *Final report. Caries diagnosis and experimental caries conference.* Columbus, Ohio, Ohio State University Research Foundation, pp 201—209

Raper, H. R. 1931. Radiodontia in operative dentistry. J. Amer. Dent. Ass. 18, 716—732

Richardson, A. S., Hole, L. & Williams, June F. 1971. The role of radiographs in clinical studies of caries-inhibitory agents. J. Pub. Health. Dent. 31, 158—165

Rugg-Gunn, A. J. 1972. Reliability and partial recording in caries incremental studies. Ph. D. Thesis, University of Manchester, p. 80

Slack, G. L., Berman, D. S., Martin, W. J. & Young, J. 1967. Clinical testing of a stannous fluoride-insoluble metaphosphate dentifrice in Kent school girls. Brit. Dent. J. 123, 9—16

Slack, G. L., Berman, D. S., Martin, W. J. & Hardie, J. M. 1967. Clinical testing of a stannous fluoride-calcium pyrophosphate dentifrice in Essex school girls. Brit. Dent. J. 123, 26—33

Torell, P. 1967. Reversals in caries diagnosis. Acta Odont. Scand. 25, 191—203

Walker, A. 1931. Detection of initial enamel lesions. Dent. Radiogr. & Photogr. 4, 3—5

Wittich, H. C. 1950. The effect of topical application of sodium fluoride upon deciduous teeth. Northw. Dent. 29, 113—114