A comparison of radiographic occlusal and approximal caries diagnoses made by 240 dentists

Ivar Espelid and Anne Bjørg Tveit
Faculty of Dentistry, University of Bergen, Bergen, Norway

Espelid I, Tveit AB. A comparison of radiographic occlusal and approximal caries diagnoses made by 240 dentists. Acta Odontol Scand 2001;59:285–289. Oslo. ISSN 0001-6357.

Radiographs of occlusal ($n = 20$) and approximal ($n = 24$) surfaces of extracted teeth were examined by 240 dentists before participating in continuing education courses dealing with caries diagnosis and treatment decisions. The radiographic caries diagnoses were treated in accordance with the receiver operating characteristic (ROC) technique, in which the area beneath the ROC curve (A_z value) indicates the quality of the observations. The frequencies of false positives made in dentin radiographically were higher for approximal (20.7%) than for occlusal caries (12.3%). The quality of pooled radiographic diagnoses of occlusal dentin lesions for all observers was significantly better than diagnoses of approximal caries in dentin. A statistically significant relationship between the observer's qualities of diagnosis of caries on approximal and occlusal surfaces ($P = 0.045$) was found. For diagnosis of dentin caries on approximal surfaces the mean Cohen kappa was 0.74 (standard deviation (*s*), 0.12; range, 0.39–0.95), and the corresponding values for occlusal surfaces were 0.70 (*s*, 0.14; range, 0.25–0.98). In the material under study the dentists were at least as good at diagnosing dentin caries occlusally as approximally. To avoid overtreatment, the observer's diagnostic threshold should ideally be adjusted towards strict criteria when a positive diagnosis is synonymous with a filling. The diagnostic thresholds were stricter in diagnosing occlusal surfaces than for approximal surfaces, indicating a more optimal strategy among dentists while diagnosing occlusal dentin lesions in a population with low caries prevalence. ☐ *Curve, receiver operating characteristic; diagnostic errors; kappa statistics; observer variation*

Ivar Espelid, Faculty of Dentistry, Department of Odontology–Pedodontics, University of Bergen, Årstadveien 17, NO-5009 Bergen, Norway. Tel: +47 55586573, e-mail: ivar.espelid@odont.uib.no

The benefit of the radiographic information in the diagnosis of approximal caries is undoubted in modern dentistry (1). For occlusal caries diagnosis an increasing number of papers focus on the advantage when supplementary information from the radiograph is utilized (2–16). Richardson & McIntyre (11) found an additional diagnostic yield from radiographs, at the dentin level of diagnosis, of 201% for occlusal surfaces, which was almost the same as for approximal surfaces, with the clinical examination detecting only 33%–34% of the total amount of caries diagnosed. However, a questionnaire study showed that Scottish dentists considered radiography to be of less value for detecting occlusal caries than approximal caries (17). This might indicate that the easy clinical access to the occlusal surface often leads to a diagnosis based on clinical findings only.

The outcome of a diagnostic process like caries diagnosis is in practice based on the observer's judgement and interpretation of the findings. Experience in reading radiographic images differs among clinicians with regard to detection of approximal and occlusal caries (15, 18, 19). The awareness and the experience of 'hidden caries' on occlusal surfaces—that is, a carious lesion seen in dentin on a radiograph where clinically the occlusal enamel appears sound or only minimally demineralized—has made clinicians focus more on the value of radiographic information (20). Our purpose was to ascertain whether there is a difference in the observer performance and interobserver variation when dentists read radiographs of occlusal and approximal surfaces.

Materials and methods

Extracted molars and premolars that had been stored in 2% benzalkonium chloride since extraction were used. The history of the teeth was unknown, but most of the molars were third molars, and most of the premolars were young teeth that apparently had been extracted for orthodontic reasons. The teeth were cleaned with pumice in water, thoroughly dried with air, and examined clinically (visually and careful probing). After teeth with caries and non-caries defects on buccal and lingual surfaces had been excluded, 20 teeth were selected for the occlusal diagnosis, whereas 24 teeth were selected for the approximal diagnosis group. Only one approximal surface from each tooth was included. Of 24 approximal surfaces 22 belonged to premolars and 2 to molars, and of 20 occlusal surfaces the corresponding numbers were 7 and 13. The teeth were radiographed tangentially to the approximal surfaces as in a bitewing projection. Exposures were made at 65 kV and 15 mA, using a dental X-ray machine with an electronic timer (Ritter II, Explorer); 1.2-sec exposure time, a focus–object distance of 32.0 cm, and an object–film distance of 1.5 cm were used. Close to the object on the focus side a 1-cm-wide water-filled plastic container was placed to simulate soft tissues. Kodak Ultraspeed DF-58 (double pack) was used. All radiographs were automatically processed, fixed, and rinsed by means of a standardized procedure in the same batch. The radiographs were duplicated on Kodak X-Omat film. This film may increase the original contrast (21).

To record the 'true' extent of the occlusal lesions toward the pulp, the buccal part of the teeth was ground away with a cylinder-shaped diamond bur in a water-cooled air rotor. When approaching the carious lesion, the tooth substance was very carefully removed, to record the size of the lesion in relation to the total enamel–dentin thickness. Any visible color change of enamel and dentin and/or reduced hardness compared with normal tissue detected by probing and/or cavity formation was taken as evidence of caries. The approximal lesions were examined during subsequent drilling from the occlusal aspect in the same manner as described for the occlusal caries validation. Eight lesions were found to penetrate into dentin, eight were confined to enamel, and eight surfaces were sound in the approximal group. The corresponding values in the occlusal group were 10, 3, and 7. Three approximal and seven occlusal dentin lesions were confined to the outer third of dentin. We validated the teeth with the naked eye during the drilling procedure and made a joint decision after discussion of each case. This method was chosen because the validation should be as close to the clinical procedure as possible.

In connection with continuing education courses dealing with caries diagnosis and treatment decisions, pre-tests were conducted. The pre-tests involved examination of the radiographs by 240 dentists in Norway. The reply rate differed from course to course but was on average about 75%.

For each of the 44 surfaces and for each area (enamel, outer third of dentin, and inner two thirds of dentin) the clinicians were asked to give their opinion on whether caries was present or not on the basis of the following grading: score 1 = caries not present; score 2 = caries probably not present; score 3 = equal likelihood that caries is present and not present; score 4 = caries probably present; and score 5 = caries present.

The diagnoses were treated in accordance with the receiver operating characteristic (ROC) technique, in which the area beneath the ROC-curve (A_z) indicates the quality of the observer performance, which is a function of the radiographs, viewing conditions, and the observer. For comparison of observer performance (diagnostic quality) the observers' registrations were pooled within each group on the basis of lesion severity. The A_z values were compared by means of non-parametric statistics (Wilcoxon), as described by Hanley & McNeil (22). When caries scores were dichotomized, the coding was as follows: scores 1, 2, or 3 = caries not present; scores 4 or 5 = caries present. The *t* test was used for comparison of means. Statistical significance was set at 5%.

To what extent observers agreed that dentin caries was radiographically present in the outer third of dentin was measured by means of linear weighted Cohen kappa (23, 24). This measure of agreement (κ) is corrected for agreement by chance. The weight gave full credit when the caries diagnoses were identical, some credit when the diagnoses differed by one grade (67%), and 33% when they differed by two grades. κ was calculated for each observer by comparing with a standard, which was the majority score for each surface. Four approximal enamel lesions were excluded in these analyses to obtain equal groups with 20 teeth in each, which is required in κ statistics.

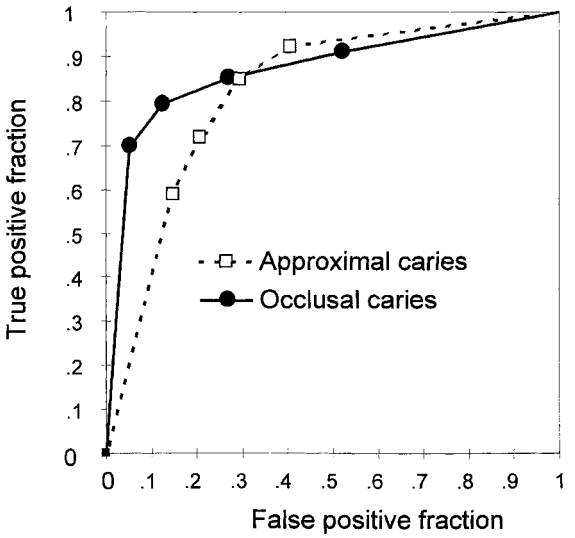

Fig. 1. Receiver operating curves (ROC) showing observer performance for 240 observers (pooled) who diagnosed dentin caries radiographically. The areas beneath ROC curves (A_z) were 0.82 (standard error, 0.006) for approximal caries and 0.87 (standard error, 0.006) for occlusal caries, indicating that the mean quality of diagnoses was a little better for the occlusal caries group ($P < 0.001$).

Results

The ROC curve (Fig. 1) gives a graphic representation of the true positive fraction (sensitivity) and the false positive fraction. The closer to the upper left corner the curve is located, the better is the observer's performance. The area beneath the curve (A_z value) gives a numerical expression of observer's performance (quality). Diagnosis by chance will give an area of 0.5, whereas perfect performance gives 1. Normally, values will be between those outer limits. The quality of dentin lesions diagnosed in dentin radiographically was best for occlusal caries (Fig. 1). It is obvious from Fig. 1 that if the pooled scores 4 + 5 were considered the diagnostic threshold, the fraction of false positives (FP) in dentin was considerably higher for approximal surfaces (20.7%) than for occlusal ones (12.3%).

Correlations between the areas beneath ROC curves of approximal diagnoses and occlusal diagnoses of dentin caries are shown in Fig. 2. The regression analysis includes ROC areas from 89 of 240 observers. The computer program for calculating areas beneath the ROC curves required that most of the response categories were used, and only 89 dentists fulfilled this criterion for both occlusal and approximal caries. The mean for omitted ROC areas

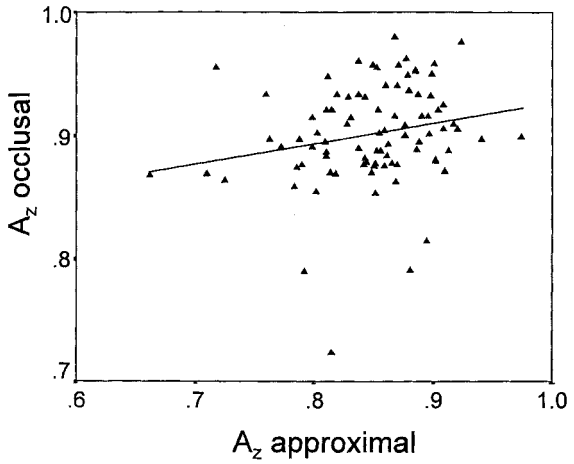

Fig. 2. Scatter plot with the area beneath receiver operating curves curves (A_z) representing radiographic diagnosis in dentin. The occlusal and approximal A_z values that are plotted were entered into a regression analysis ($R^2 = 0.045$, $y = 0.761 + 0.165x$; $P < 0.001$).

due to incomplete pairs was tested to ascertain whether they differed from the areas that were included in the regression analyses. There were no significant differences between the omitted ($n = 33$) and included ($n = 89$) 'occlusal' curves ($P = 0.55$) or between the omitted ($n = 78$) and the included ($n = 89$) 'approximal' curves ($P = 0.25$). Although the plots are widely spread along the regression line, there is a statistically significant relationship in the diagnostic quality (A_z) in the two different locations ($P = 0.045$).

The fraction of false positives made when sound surfaces or surfaces with enamel lesions were examined in the outer dentin radiographically ranged from 0 to 0.700 (occlusal surface) to 0 to 0.625 (approximal surface) among the observers. Table 1 shows corresponding cross-tabulation of the number of false positives for each dentist on the basis of the type of surface diagnosed. When these data were put into a regression analysis, there was a small but statistically significant correlation between FP occlusal and FP approximal ($r^2 = 14.1$, $P < 0.001$). The correlation coefficient (Pearson) was 0.376 ($P < 0.01$).

The mean κ value was 0.74 (s, 0.12; range, 0.39–0.95) for the approximal surfaces and 0.70 (s, 0.14; range, 0.25–0.98) for the occlusal surfaces. The difference between the means was statistically significant ($P = 0.0006$). Of 240 observers, 138 had their highest κ value when diagnosing approximally, whereas 102 had their highest κ diagnosing occlusally.

Discussion

The observers' performance when examining the radiographs depends on skill, experience, viewing conditions, and the material under examination (25). This material was the same for each observer, and the other variables were most likely kept constant for each observer. The area beneath the ROC curve was used as a relative estimate of the quality of the diagnoses, which is not affected by the strategy chosen by the observer (26). The material was selected so that each group of teeth contained lesions of similar severity. Thereby, the results could illustrate whether dentists consequently scored better on one type of surface than on another, presuming that the lesions were of comparable size.

The present results do not indicate that dentists are poorer in diagnosing dentin caries in the occlusal surfaces than in the approximal ones, even though many dentists have expressed doubts about the value of the radiograph in occlusal caries diagnosis (17).

The radiographic picture of caries differs in accordance with the location. The approximal surface is more or less curved, and loss of minerals in the outer part of the enamel is less hidden by sound tooth structure than for occlusal caries in the two-dimensional representation of the three-dimensional tooth. In addition, the occlusal surface is

Table 1. Cross-table showing distribution of false-positive diagnoses made by 240 dentists in the outer third of dentin in occlusal and approximal surfaces. All surfaces were validated as sound or as having a lesion confined to enamel (10 occlusal and 16 approximal surfaces)

	Occlusally									
	0	1	2	3	4	5	6	7	Total	Cumulated percentage
Approximally										
0	4								4	1.7
1	11	9	1						21	10.4
2	15	20	3	3					41	27.5
3	33	23	15	5	3	1			80	60.8
4	16	15	8	4	2	1	1	1	48	80.8
5	3	10	7	7	1	1			29	92.9
6		3	4	2				1	10	97.1
7		2	2		1				5	99.2
10			1		1				2	100.0
Total	82	82	41	21	8	3	1	2	240	
Cumulated percentage	34.2	68.3	85.4	94.2	97.5	98.8	99.2	100.0		

vulnerable to 'noise' created by the occlusal anatomy. There is a tendency to make false-positive scores just beneath the dentinoenamel junction, probably due to the Mach-band effect (15). Therefore, it was difficult to match the groups exactly when composing the groups of teeth in the study. However, the total number of dentin lesions in the two groups (10 in the occlusal group and 8 in the approximal group) was very close. There were five deep dentin lesions in the approximal group, compared with three in the occlusal group, which is in favor of higher detectability of the approximal dentin lesion. Taking this fact into account strengthens rather than weakens the finding—namely, the slightly better diagnostic quality registered for the occlusal dentin lesions.

On the basis of the numbers in Table 1 it can be calculated that 50 dentists (21%) have more than 4 false positives in either the occlusal, approximal, or both surfaces when the outer third of the dentin is examined radiographically. Of a total of 1089 false-positive decisions these dentists accounted for 392 (36%). However, only four dentists (2%) did not have any false positives at all. A comparison of occlusal and approximal diagnoses indicates that overscoring is far more common in approximal surfaces. In a clinical situation the proportion of false positives for the occlusal caries will probably be reduced more than for approximal lesions, since the occlusal surface always is available for clinical examination.

The quality of diagnosis in the two actual locations should be interpreted with this in mind, and the major intention of the study was to disclose any significant trend with regard to the quality of diagnoses among the participants. The kappa value in this study may also be read as an index of quality provided the majority opinion was the gold standard. The interobserver agreement, comparing each observer with a 'mean' score for the whole group, was higher for approximal caries (0.74) than for occlusal caries (0.70), but both values were within the range of substantial agreement. The variation did not differ substantially between occlusal and approximal diagnoses but was lowest in the latter group. An explanation might be that the observers are more accustomed to radiographic diagnosis on the approximal surface. Some dentists have the opinion that radiography is of less value for detecting occlusal caries than approximal lesions (17). According to Landis & Koch (27), a score of 1.00–0.81 indicates almost perfect agreement, 0.80–0.61 indicates substantial agreement, and 0.60–0.41 moderate agreement. The present values exceeded those reported by Lavonius et al. (28) for visual occlusal caries diagnoses on extracted teeth. With regard to interobserver agreement they found a κ value of 0.46 before and 0.59 after calibration, but the results cannot be directly compared with those in the present study owing to the different validation methods.

The cut-off points on the respective ROC curves reflect the observer strategy (Fig. 1). The sensitivity and specificity values for the combined confidence scores 4 + 5 of diagnosis in the dentin of dentin lesions (high confidence that caries was present) are 79% and 87% for occlusal caries, respectively, and the corresponding values for approximal caries are 72% and 79%. The binary outcome of a diagnostic test is also influenced by the observers' conscious or non-conscious evaluation of the relative costs associated with classifying, for example, a tooth surface with caries as sound versus classifying a sound surface as carious (29). If the present strategies were used clinically, they would have led to a higher frequency of false-positive diagnoses of the approximal surfaces. However, the results should be interpreted with care owing to the non-clinical setting, the low number of surfaces used, and the fact that the ROC curves cross each other.

It has been stated that the education and experience level of the examiners, type of practice, and patient characteristics influence the ability to diagnose occlusal caries (30, 31). In the present study the 240 dentists represented different types of clinicians—both dentists from public dental health service and private practitioners.

From a clinical point of view it is of particular interest to focus on diagnosis of dentin lesions because these may lead to decisions to make a restoration. Table 1 shows that most dentists tend to overscore lesions (false-positive diagnoses) in the outermost part of dentin and that there is a relationship between the false-positive fractions approximally and occlusally. Downer & Moles (32) have made an interesting computer simulation of the influence over time of relevant factors on health gain from restorative dental treatment. They include sensitivity and specificity of the diagnoses on the basis of bitewing radiography and show that variation in these factors plays an important role for health gain when utility values for different outcomes are included in the model. The authors conclude that caution is advisable when decisions to restore are based on bitewing radiography. The great disparity among the 240 dentists in the present study indicates that there is a need for reducing the variation. It has been shown that education and clinical experience increase interexaminer reproducibility in radiographic diagnoses of occlusal caries (33), and for clinical diagnoses of occlusal caries the interexaminer reproducibility can be improved with a calibration process (28).

Conventional radiography is still a significant and reasonably accurate diagnostic method for the detection of approximal caries in enamel and dentin (15, 25, 34). However, occlusal caries needs to be in dentin for radiographic detection (35), and still there is tendency for false positives in the outermost part of dentin (15). Ricketts et al. (36) claim that 'although the bite-wing radiograph may be regarded as a safety net for the diagnosis of occlusal caries, it must be interpreted with caution bearing in mind the possibility of false positive diagnoses'.

The present results indicate that clinicians showing high quality in diagnosing approximal caries are likely to show high quality in occlusal caries diagnoses as well. The false positives in dentin are of particular interest, since most dentists in Norway report that a positive caries diagnosis in

dentin might be a criterion for instigating operative treatment (37). In principle, there is a positive relationship between false positives and true positives. To avoid overtreatment, the observer's diagnostic threshold should ideally be adjusted towards strict criteria when a positive diagnosis will lead to the placing of a filling. The diagnostic thresholds were stricter in diagnosis of occlusal surfaces than approximal surfaces, indicating a more optimal strategy among dentists while diagnosing in dentin of occlusal surfaces in a population with low caries prevalence.

References

1. Kidd EAM, Pitts NB. A reappraisal of the value of the bitewing radiograph in the diagnosis of posterior approximal caries. Br Dent J 1990;169:195–200.
2. Shearer AC, Wilson NH. Radiography for the diagnosis of dental caries in posterior teeth. Clin Prev Dent 1989;11:24–8.
3. Weerheijm KL, van Amerongen WE, Eggink CO. The clinical diagnosis of occlusal caries: a problem. J Dent Child 1989;56: 196–200.
4. Wenzel A, Fejerskov O, Kidd EAM, Joyston-Bechal S, Groeneveld A. Depth of occlusal caries assessed clinically, by conventional film radiographs, and by digitized, processed radiographs. Caries Res 1990;24:327–33.
5. Wenzel A, Hintze H, Mikkelsen L, Mouyen F. Radiographic detection of occlusal caries in noncavitated teeth. A comparison of conventional film radiographs, digitized film radiographs, and Radio VisioGraphy. Oral Surg Oral Med Oral Pathol 1991;72: 621–6.
6. Wenzel A, Larsen MJ, Fejerskov O. Detection of occlusal caries without cavitation by visual inspection, film radiographs, xeroradiographs, and digitized radiographs. Caries Res 1991;25: 365–71.
7. Russell M, Pitts NB. Occlusal caries diagnosis: radiovisiography versus bite-wing radiography. Caries Res 1991;25:217.
8. van Amerongen JP, Penning C, Kidd EAM, ten Cate JM. An in vitro assessment of the extent of caries under small occlusal cavities. Caries Res 1992;26:89–93.
9. Weerheijm KL, Groen HJ, Bast AJJ, Kieft JA, Eijkman MAJ, van Amerongen WE. Clinically undetected occlusal dentine caries: A radiographic comparison. Caries Res 1992;26:305–9.
10. Nytun RB, Raadal M, Espelid I. Diagnosis of dentin involvement in occlusal caries based on visual and radiographic examination of the teeth. Scand J Dent Res 1992;100:144–8.
11. Richardson PS, McIntyre IG. The difference between clinical and bitewing detection of approximal and occlusal caries in Royal Air Force recruits. Community Dent Health 1996;13:65–9.
12. Poorterman JH, Weerheijm KL, Groen HJ, Kalsbeek H. Clinical and radiographic judgement of occlusal caries in adolescents. Eur J Oral Sci 2000;108:93–8.
13. Ketley CE, Holt RD. Visual and radiographic diagnosis of occlusal caries in first permanent molars and in second primary molars. Br Dent J 1993;174:364–70.
14. Kidd EAM, Ricketts DNJ, Pitts NB. Occlusal caries diagnosis: a changing challenge for clinicians and epidemiologists. J Dent 1993;21:323–31.
15. Espelid I, Tveit AB, Fjelltveit AB. Variations among dentists in radiographic detection of occlusal caries. Caries Res 1994;28: 169–75.
16. Waraporn T, Tveit AB, Espelid I. Occlusal caries-variations in diagnosis and treatment decisions among university teachers in Norway and Thailand. Eur J Paed Dent 2000;4:161–9.
17. Pitts NB, Fyffe HE. Scottish dentists' use of and opinions regarding bitewing radiography. Dentomaxillofac Radiol 1991; 19:214–8.
18. Syriopoulos K, Sanderink GC, Velders XL, van der Stelt PF. Radiographic detection of approximal caries: a comparison of dental films and digital imaging systems. Dentomaxillofac Radiol 2000;29:312–8.
19. Mileman PA, Mulder E, van der Weele L. Factors influencing the likelihood of successful decisions to treat dentin caries from bitewing radiographs. Community Dent Oral Epidemiol 1992; 20:175–80.
20. Weerheijm KL. Occlusal 'hidden caries'. Dent Update 1997;24: 182–4.
21. Price C. A densitometric evaluation of two radiographic duplicating films under differing conditions of exposure and processing. Oral Surg Oral Med Oral Pathol 1980;50:190–4.
22. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143:29–36.
23. Amarante E, Raadal M, Espelid I. Impact of diagnostic criteria on the prevalence of dental caries in Norwegian children aged 5, 12 and 18 years. Community Dent Oral Epidemiol 1998;26:87–94.
24. Bulman JS, Osborn JF. Statistics in dentistry. London: British Dental Association; 1989.
25. Espelid I. Radiographic diagnoses and treatment decisions on approximal caries [thesis]. Bergen: University of Bergen; 1987.
26. Swets JA, Pickett RM. Evaluation of diagnostic systems. Methods from signal detection theory. New York: Academic Press; 1982.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
28. Lavonius E, Kerosuo E, Kallio P, Pietila I, Mjör IA. Occlusal restorative decisions based on visual inspection—calibration and comparison of different methods. Community Dent Oral Epidemiol 1997;25:156–9.
29. McNeil BJ, Keller E, Adelstein SJ. Primer on certain elements of medical decision making. N Engl J Med 1975;293:211–5.
30. Kay EJ, Watts A, Paterson RC, Blinkhorn AS. Preliminary investigation into the validity of dentists' decisions to restore occlusal surfaces of permanent teeth. Community Dent Oral Epidemiol 1988;16:91–4.
31. Bader JD, Shugars DA. Understanding dentists' restorative treatment decisions. J Public Health Dent 1992;52:102–10.
32. Downer MC, Moles DR. Health gain from restorative dental treatment evaluated by computer simulation. Community Dent Health 1998;15:32–9.
33. Lazarchik DA, Firestone AR, Heaven TJ, Filler SJ, Lussi A. Radiographic evaluation of occlusal caries: effect of training and experience. Caries Res 1995;29:355–8.
34. Wenzel A, Pitts N, Verdonschot EH, Kalsbeek H. Developments in radiographic caries diagnosis. J Dent 1993;21:131–40.
35. Pine CM, ten Bosch JJ. Dynamics of and diagnostic methods for detecting small carious lesions. Caries Res 1996;30:381–8.
36. Ricketts DNJ, Kidd EAM, Smith BGN, Wilson RF. Clinical and radiographic diagnosis of occlusal caries: a study in vitro. J Oral Rehabil 1995;22:15–20.
37. Tveit AB, Espelid I, Skodje F. Restorative treatment decisions on approximal caries in Norway. Int Dent J 1999;49:165–72.

Received for publication 10 November 2000
Accepted 11 April 2001