DOI: 10.1111/j.1365-263X.2006.00807.x

Use of a ranked scoring system to detect occlusal caries in primary molars

SUZANNE DUNKLEY & PAUL ASHLEY

Unit of Paediatric Dentistry, Eastman Dental Institute, University College London, London, UK

International Journal of Paediatric Dentistry 2007; 17: 267–273

Objective. The available literature concerning systems for occlusal caries detection in primary teeth is limited. The aim of this study was to compare *in vitro* a ranked scoring system for occlusal caries detection in primary molar teeth with bitewing radiography and to investigate the most accurate cut-off point for dentine caries detection when using this ranked scoring system.

Methods. Sixty primary molar teeth were examined *in vitro* visually and by bitewing radiography (two examiners) using Ekstrand' criteria.

Results. Histological examination revealed that 13% were sound, 38% had enamel caries and the remainder had dentine caries. Similar accuracy was seen when comparing bitewing radiographs and visual examination for caries detection at the d3 threshold, while visual examination was more accurate at the d1 threshold. The most suitable cut-off for diagnosing caries at the d1 threshold was V1 (no/slight change in enamel translucency after air drying). Either V2 (opacity/discoloration visible without air drying) or V3 (enamel breakdown in opaque/discoloured enamel and/or greyish discoloration) were suitable cut-off points at the d3 threshold.

Conclusion. Visual examination was more accurate than bitewing radiographs for detection at the d1 threshold (the cut-off point of V1). Either V2 or V3 can be used for caries detection at the d3 threshold.

Introduction

The available literature concerning systems for the detection of occlusal caries in primary teeth is limited. A systematic review of the literature conducted by the Agency for Healthcare Research and Quality of the United States Department of Health and Human Services[2] found a 'virtual absence of any assessment of diagnostic methods applied to primary teeth', the report went on to state that 'too few assessments addressed diagnosis on primary teeth to permit conclusions to be drawn' on the validity of the existing diagnostic systems.

Since then, several studies have been published looking at occlusal caries detection in primary teeth and these have largely focused on DIAGNOdent (Kavo, Biberach, Germany). Reported results to date[3–6] have shown that DIAGNOdent does have utility as a diagnostic tool, although use of visual inspection may offer similar accuracy (where accuracy can be defined as the extent to which the test results reflect true disease status). Only two of these studies[3,6] used ranked criteria for visual detection. The available evidence from permanent teeth[1] suggests that use of ranked criteria for visual detection will boost diagnostic yield. Both of these studies, however, used different cut-off points for the detection of enamel or dentine caries.

Further work is required in order to determine the most accurate way of applying these criteria for occlusal caries detection in primary teeth and to compare this method of caries detection to bitewing radiography.

Therefore, the aims of this study were to compare *in vitro* the use of a ranked scoring system for the detection of occlusal caries in primary molar teeth with bitewing radiography, and to investigate the most accurate cut-off point for dentine caries detection in primary molar teeth when using this ranked scoring system.

Correspondence to:
Paul Ashley, Unit of Paediatric Dentistry, Eastman Dental Hospital, 256 Gray's Inn Road, London WC1X 8LD, UK.
E-mail: p.ashley@eastman.ucl.ac.uk

Declaration of interest: No conflicts of interest

Materials and methods

Ethics approval was obtained from Eastman Dental Hospital Joint Research and Ethics Committee. Recruitment was done prospectively from patients attending an outpatient, extraction-only, general anaesthetic clinic at the Eastman Dental Hospital, London, UK.

The teeth were stored in 10% buffered formalin immediately following extraction and had none of the following:

- occlusal restorations;
- fissure sealants;
- developmental defects; and
- frank cavitation, i.e. cavitation visible on initial examination.

The occlusal surfaces were cleaned for 60 s with a bristle brush, and a pumice and water slurry. The teeth were then rinsed with sterile water and coded.

Bitewing radiographs were taken of the teeth. A jig was constructed so that four teeth could be radiographed per standard intraoral film (Kodak F-speed, Kodak Ltd, Hemel Hempstead, UK). A 10-mm-thick sheet of Plexiglass was placed 3 cm from the film to mimic the buccal soft tissues. The X-ray machine (Densomat, Philips, Eindhoven, the Netherlands) was set at 65 kV and 7.5 mA, with a focal point to film/imaging plate distance of 30 cm. The exposure time was 0.22 s. The teeth were all radiographed on the same day and the films were developed immediately following exposure. An automatic processor (Velopex, Medivance Instruments, London, UK) was used to develop the films.

The teeth were then examined by two examiners (P.A. and S.D.) for the presence or absence of occlusal caries using the criteria described by Ekstrand *et al.*[1] (Tables 1 & 2). One examiner (P.A.) was experienced in the detection of caries using a visual diagnostic system and demonstrated the use of such a system with 10 primary molars to the second examiner (S.D.). These teeth did not form part of the main study set. Both examiners were experienced in the use of the radiographic diagnostic system. Prior to the examinations, a third party randomly recoded the teeth. The two examiners independently examined the teeth in random order, thus being blind to one another's results.

Table 1. Ekstrand radiographic criteria.*

Rating	Criterion
R0	No radiolucency visible in the enamel
R1	Radiolucency visible in the enamel
R2	Radiolucency visible in the dentine, but restricted to the outer third of the dentine
R3	Radiolucency extending to the middle third of the dentine
R4	Radiolucency in the pulpal third of the dentine

*Radiographs were viewed under standardized conditions using a light box and × 1.5 magnification.

Table 2. Ekstrand visual criteria.*

Rating	Criterion
V0	No or slight change in enamel translucency after prolonged air drying (> 5 s)
V1	Opacity or discoloration hardly visible on the wet surface, but distinctly visible after air drying
V2	Opacity or discoloration distinctly visible without air drying
V3	Localized enamel breakdown in opaque or discoloured enamel and/or greyish discoloration from the underlying dentine
V4	Cavitation in opaque or discoloured enamel exposing the dentine

*The tooth under examination was dried for 10 s using compressed air and examined under a standard dental operating light.

Reproducibility of visual and radiographic examination

Inter- and intra-examiner repeatability was determined by re-examining a randomly selected subgroup of 25% of the original sample ($n = 15$ teeth). The examiners undertook the second examinations independently and more than 4 days after the original examination. The examiners were blind to the original results.

Histological validation

Teeth were hemi-sectioned in a mesial–distal direction through the fissure pattern with a high-speed drill and fine diamond bur. Wet sections were viewed under a microscope at × 10 magnification by the principal examiner (S.D.). The presence or absence of occlusal caries was recorded using the following criteria (Table 3). The criteria chosen were not those used in the Ekstrand study[1], but were selected to facilitate comparison with other studies[3,5].

Table 3. Histological criteria.

Rating	Criterion
H0	No enamel demineralization
H1	Enamel demineralization
H2	Demineralization involving the outer half of the dentine
H3	Demineralization involving the inner half of the dentine

The examiner was blind to the radiographic diagnostic test results. The reproducibility of the histological validation was determined by re-examining the sections of a subgroup of 15 randomly selected teeth. The examiner was blind to the previous results.

Statistical methodology

Data obtained were entered onto the SPSS, Version 12.01, computer program (SPSS Inc., Chicago, IL, USA). Receiver operating characteristic (ROC) analysis was conducted to compare the diagnostic performance of visual inspection versus bitewing radiography at the d1 threshold (sound versus enamel/dentine) and at the d3 threshold (sound/enamel versus dentine caries). Areas were compared using the method described by Hanley and McNeil[7]. In addition, likelihood ratios (where the likelihood ratio is defined as the ratio of the probability of the specific test result in people who do have the disease to the probability in the people who do not) were calculated for the visual system using each point on the ordinal classification scale (i.e. V1, V2, etc.) as a cut-off point for the detection of occlusal caries at the two histological cut-offs used above. This additional data was used to help determine the 'best' cut-off point.

Data from the repeatability exercise were compared using the weighted kappa statistic[8]. The intra-examiner repeatability was calculated using a subgroup of 25% of the original sample ($n = 15$). The interexaminer repeatability was calculated using the whole sample.

Results

Sixty teeth, 35 first and 25 second primary molars, were collected after extraction at an outpatient general anaesthetic clinic at the Eastman Dental Hospital. Histological examination of the 60 teeth, revealed eight (13%) teeth to be sound, 23 (38%) with enamel caries and 29 (48%) teeth with dentine caries (17 with lesions involving the inner half of the dentine).

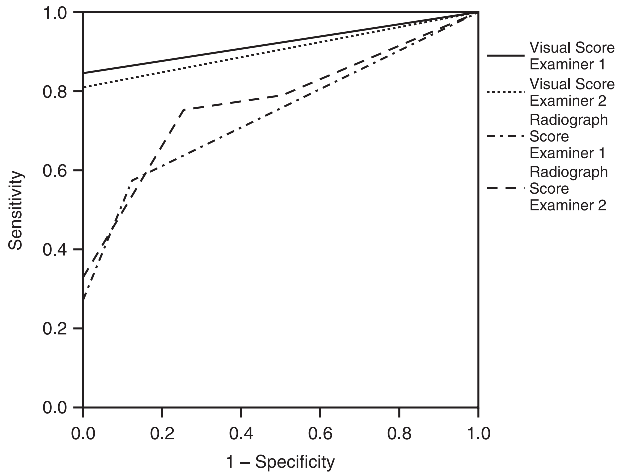

Fig. 1. Receiver operating characteristic analysis for both examiners, and the visual and radiographic systems for the detection of caries at the d1 threshold.

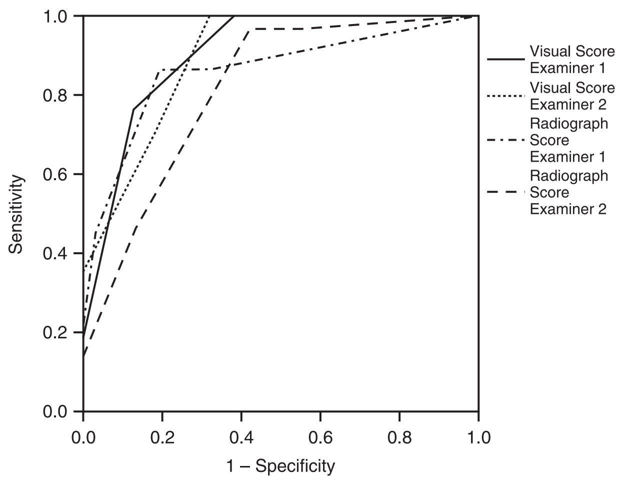

Fig. 2. Receiver operating characteristic analysis for both examiners, and the visual and radiographic systems for the detection of caries at the d3 threshold.

Receiver operating characteristic curves showing the overall performance of each system for each examiner at the two histological levels (d1 and d3) are shown in Figures 1 and 2. There were no statistically significant differences ($P > 0.05$) between examiners, or

Table 4. Likelihood ratios of visual classifications for the detection of occlusal caries at the d1 threshold (95% confidence intervals in parentheses).

Rating	Likelihood ratio positive test		Likelihood ratio negative test	
	Examiner 1	Examiner 2	Examiner 1	Examiner 2
V1	*	*	0.15 (0.08, 0.29)	0.19 (0.11, 0.34)
V2	*	*	0.12 (0.11, 0.34)	0.25 (0.16, 0.40)
V3	*	*	0.50 (0.38, 0.66)	0.48 (0.36, 0.64)
V4	*	*	0.90 (0.83, 0.99)	0.81 (0.71, 0.92)

*Values for the positive likelihood ratio could not be calculated.

Table 5. Likelihood ratios of visual classifications for the detection of occlusal caries at the d3 threshold (95% confidence intervals in parentheses).

Rating	Likelihood ratio positive test		Likelihood ratio negative test	
	Examiner 1	Examiner 2	Examiner 1	Examiner 2
V1	2.08 (1.44, 2.97)	2.40 (1.58, 3.61)	0	0
V2	2.60 (1.70, 4.00)	3.10 (1.86, 5.16)	0	0
V3	5.88 (2.3, 15.01)	3.74 (1.76, 7.94)	0.28 (0.14, 0.54)	0.34 (0.19, 0.63)
V4	*	*	0.83 (0.70, 0.98)	0.66 (0.50, 0.85)

*Values for the positive likelihood ratio could not be calculated.

between visual or radiographic detection for the examiners at the d1 threshold [areas under the ROC curve (A_z) compared]. There was a significant difference ($P < 0.05$) between visual and radiographic detection for examiner 1 (in favour of visual detection) at this threshold. There were no statistically significant differences ($P > 0.05$) between examiners, or between visual or radiographic detection for the detection of caries at the d3 threshold when comparing A_z.

Likelihood ratios for the visual system using each point on the ordinal classification scale as a cut-off point for the detection of occlusal caries at the two histological cut-offs used are summarized in Tables 4 and 5. When diagnosing caries at the d3 threshold, values for the likelihood ratio of a positive test increased with increase in rank, indicating that, the higher the rank used as a cut-off point, the more likely the patient was to have the disease. The same was true of values for the likelihood ratio of a negative test; however, as this is expressed as a fraction of 1, an increase indicates that, the higher the rank, the more likely a patient with a negative test was to have the disease. When diagnosing caries at the d1 threshold, the positive likelihood ratio could not be calculated because it was effectively infinity.

Table 6. Weighted kappa values of inter- and intra-examiner repeatability for the visual and radiographic diagnostic systems (95% confidence intervals in parentheses).

Examiner	Diagnostic system	
	Visual	Radiographic
1	0.93 (0.86–0.99)	1.00
2	0.94 (0.88–0.99)	0.87 (0.70–1.00)
1 versus 2	0.93 (0.90–0.97)	0.75 (0.63–0.88)

The results for the inter- and intra-examiner repeatability are summarized in Table 6. Examiner 2 gave the best repeatability for the visual detection (0.94) and examiner 1 the best repeatability for the radiographic diagnostic system (1.00). The visual diagnostic system had the highest kappa value (0.93) for interexaminer repeatability. The kappa values showed that there was excellent repeatability for both diagnostic systems both between the examiners and by the same examiner.

Fifteen of the histological sections were reassessed and 14 were scored as previously; however, one that was originally scored as having caries into the outer half of the dentine was subsequently scored as having caries in the inner half of the dentine.

Discussion

The sample size ($n = 60$) in this study was comparable to previous studies[3,5]. The distribution of caries lesions within the sample of teeth represented a good range of lesion depths, with 13% of the sample being sound, 38% having enamel lesions and 48% having dentine lesions on histological examination. The sample contained a sizeable number of teeth with pre-cavitated lesions, and this enabled the ability of the diagnostic systems to detect early lesions to be assessed.

Data from the ROC curve analysis suggests that there was little to choose between the radiographic and visual systems when detecting at the d3 threshold; however, there were differences when considering detection at the d1 caries threshold [examiner 1, significant difference in A_z ($P < 0.05$) in favour of visual detection]. This pattern is similar to that reported by Rocha *et al.*[6]. The poor performance of radiographs for the detection of occlusal caries at the d1 threshold is not surprising; what is interesting though is the excellent performance of the visual system, as demonstrated by the area under the ROC curve, and values for sensitivity, specificity, positive and negative predictive values.

The next aim of this study was to determine the 'best' cut-off point for the detection of caries at the d1 or d3 thresholds. This depends in part on what you intend to do with the result. If the intention is to provide a restoration, then you wish to ensure that specificity and the likelihood ratio of a positive test are both high to reduce the risk of inadvertently restoring a sound surface. This may, however, be at the cost of missing a proportion of people who require treatment. Alternatively, if the aim is to provide preventive advice, you may be more interested in a higher sensitivity and accept a lower positive likelihood ratio in order to include more people with caries at the risk of putting people without disease through your preventive programme.

If we consider the detection of any caries (d1 threshold), then the choice of cut-off is straightforward. The values for the positive likelihood ratio were 'perfect' at each cut-off point. Therefore, we want the cut-off point with the most favourable negative likelihood ratio, which is V1. Values for sensitivity and specificity were calculated for detection of d1 caries [examiner 1, sensitivity = 0.85 [95% confidence interval (CI) = 0.73, 0.92], specificity = 1 (95% CI = 0.67, 1)] at this cut-off. They were higher than those reported by Lussi and Francescut[5] (sensitivity = 0.54, specificity = 0.68) or Rocha *et al.*[6] (sensitivity = 0.82, specificity = 0.85). Lussi and Francescut[5] did not use a ranked criteria system, so this might explain the difference. Rocha *et al.*[6] did use the same system, and did report higher values for sensitivity and specificity than Lussi and Francescut, but they were not as high as those in this study.

Determining the 'best' cut-off point for detection of dentine caries (d3 threshold) is more complex. Both V1 and V4 can be discarded because of overall poor accuracy. It is more difficult to choose between V2 and V3, however. Use of V2 returns a perfect result for the likelihood ratio of a negative test (i.e. a negative result strongly suggests that there is no disease), but low values for the likelihood ratio of a positive test. Use of V3 does improve the value for the likelihood ratio of a positive test, but at the cost of a greatly increased likelihood ratio of a negative test. The decision will have to be left up to the individual operator because it is difficult to settle on a cut-off point without contextual information as to how it will be used.

How do these results compare to previous studies? As before, we can calculate values of sensitivity and specificity at individual cut-off points to facilitate comparison, in this case at V2 and V3 for the d3 threshold (calculated for examiner 1 only). At V2, sensitivity and specificity are 1.0 (95% CI = 0.88, 1.0) and 0.61 (95% CI = 0.44, 0.76), respectively; at V3 sensitivity and specificity are 0.76 (95% CI = 0.58, 0.88) and 0.87 (95% CI = 0.71, 0.95), respectively. Lussi and Francescut[5] reported values of sensitivity and specificity of 0.35 and 0.98, respectively, but they did not use ranked criteria, so this might explain the low accuracy (note the similarity to use of V4 in this study). Rocha *et al.*[6] reported values of 0.61 and 1.0 for sensitivity and specificity using V3 as the cut-off point. Whilst these values for sensitivity and specificity are higher and lower, respectively, than those reported in this study at V3, they are not that dissimilar. Attrill and Ashley[3] used the same visual criteria, and achieved a sensitivity of 0.57 for one examiner and 0.63 for a second examiner at the V2 cut-off (the corresponding values of specificity were 0.93 and 0.89). This study achieved higher values for sensitivity at V2, but lower values of specificity. Conversely, Attrill and Ashley reported low values for sensitivity (0.4) and high values for specificity at V3. This difference is difficult to account for.

The inter- and intra-examiner reproducibility was excellent for the visual diagnostic system, as expressed by weighted kappa values[8]. This suggests that the good results produced by the enhanced visual criteria can be reliably repeated, a necessary feature of any diagnostic system used in clinical practice. The weighted kappa values were also good for the radiographic diagnostic system.

This study relied on the ranked scoring system developed by Ekstrand[1]. This was developed following careful examination of occlusal carious lesions in permanent teeth and relation of their appearance to histological change. Arguably, this scale is not applicable to the primary dentition. Nevertheless, the descriptions do aid in classifying lesions, and the scale appears to show accuracy and repeatability for occlusal caries detection in primary teeth. This study also relied on histological examination of teeth as a gold standard. Conventionally, serial sections are used; however, in this study, the authors used hemi-sections since this was simpler. It could be argued that serial sectioning might have uncovered more caries; however, the available data[9] suggest that hemi-sections will detect as many lesions as serial sections when assessing occlusal surfaces.

All *in vitro* studies are of limited value since they do not reflect the difficulties encountered in the clinical setting. Therefore, the ideal study design would be an *in vivo* study, whereby the primary teeth were examined in the clinical setting, and scored according to the enhanced visual criteria prior to their extraction and subsequent histological examination. More data on the use of this approach in a primary care setting would also be useful.

In conclusion, visual examination (using the enhanced visual criteria) is as accurate as bitewing radiography for occlusal caries detection at the d3 level in primary molars, and more accurate for detection at the d1 threshold. A cut-off point of V1 is best for detection of occlusal caries at the d1 threshold. Either V2 or V3 can be used for caries detection at the d3 threshold in primary molars; the choice will be dependent on the use to which this information is put.

What this paper adds

- Visual examination can be used instead of bitewing radiography for occlusal dentine (d3) caries detection in primary molar teeth.
- Visual examination shows good accuracy for the detection of occlusal caries at the d1 threshold in primary molar teeth.

Why this paper is important to paediatric dentists?

- Early occlusal caries detection in primary molars is important because of the speed with which caries progresses in primary teeth.
- This paper helps paediatric dentists to determine when to intervene.

Acknowledgements

We would like to thank A. Johnson for help with the blinding. There was no funding for this project.

References

1 Ekstrand KR, Ricketts DNJ, Kidd EAM. Reproducibility and accuracy of three methods for assessment of demineralisation depth on the occlusal surface: an *in vitro* examination. *Caries Res* 1997; **3**: 224–231.

2 Bader JD, Shugars DA, Bonito AJ. Systematic reviews of selected dental caries diagnostic and management methods. *J Dent Educ* 2001; **65**: 960–968.

3 Attrill DC, Ashley P. Occlusal caries detection in primary teeth: a comparison of DIAGNOdent with conventional methods. *Br Dent J* 2001; **190**: 440–443.

4 Bengtson AL, Gomes AC, Mendes FM, Cichello LRD, Bengtson NG, Pinheiro SL. Influence of examiner's clinical experience in detecting occlusal caries lesions in primary teeth. *Pediat Dent* 2005; **27**: 238–243.

5 Lussi A, Francescut P. Performance of conventional and new methods for the detection of occlusal caries in deciduous teeth. *Caries Res* 2003; **37**: 2–7.

6 Rocha RO, Ardenghi TM, Oliveira LB, Rodrigues CRMD, Ciamponi AL. *In vivo* effectiveness of laser

fluorescence compared to visual inspection and radiography for the detection of occlusal caries in primary teeth. *Caries Res* 2003; **37**: 437–441.

7 Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiol* 1983; **148**: 839–843.

8 Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973; **33**: 613–619.

9 Hintze H, Wenzel A. Diagnostic outcome of methods frequently used for caries validation. *Caries Res* 2003; **37**: 115–124.