A Comparative Study of Dentists' Ability to Detect Enamel-only Proximal Caries in Bitewing Radiographs With and Without the use of AssistDent® Artificial Intelligence Software

Authors: Hugh Devlin[1] Tomos Williams[2] Jim Graham[3] Martin Ashley[4]

[1]Professor of Restorative Dentistry, Division of Dentistry, School of Medical Sciences, The University of Manchester; Director, Manchester Imaging Ltd.

[2]Honorary Research Assistant, Division of Dentistry, School of Medical Sciences, The University of Manchester; Software Manager, Manchester Imaging Ltd.

[3]Honorary Reader, Division of Informatics, Imaging and Data Sciences, School of Health Sciences, The University of Manchester; Director, Manchester Imaging Ltd.

[4]Consultant and MAHSC Honorary Professor in Restorative Dentistry and Oral Health, University Dental Hospital of Manchester, Manchester University NHS Foundation Trust.

Abstract

Enamel-only proximal caries, if detected, can be reversed by non-invasive treatments. Dental bitewing radiograph analysis is central to diagnosis and treatment planning and, when used to detect enamel-only proximal caries, it is an important tool in minimum intervention and preventive dentistry. However, the subtle patterns of enamel-only proximal caries visible in bitewing radiographs are difficult to detect and often missed by dental practitioners. This study measures the ability of dentists to detect enamel-only proximal caries in bitewing radiographs with and without the use of AssistDent® Artificial Intelligence (AI) software.

23 dentists were randomly divided into a control arm, in which no Artificial Intelligence assistance was provided, and an experimental arm in which Artificial Intelligence assistance provided on-screen prompts for potential locations of enamel-only proximal caries. All participants analysed a set of 24 bitewings, gathered from one dental hospital and 11 general dental practices, which had previously been analysed independently by a panel of 5 dento-maxillofacial radiologists and 1 professor of restorative dentistry who, between them, identified a total of 65 enamel-only carious lesions and 241 healthy proximal surfaces.

Results demonstrate that dentists using the assistive software found 75.8% of the enamel-only proximal caries compared to a 44.3% detection rate in the control group. This represents an absolute increase of 31.5% (relative increase in sensitivity of 71%). Participants in the experimental group incorrectly identified 14.6% of the healthy surfaces as having enamel-only proximal caries compared to 3.7% in the control group, an absolute increase of 10.9% (relative decrease in specificity of 11%).

T-test analysis demonstrated a statistically significant difference ($p<0.01$) between the two arms in sensitivity (true positive caries detection rate) and specificity (false positive rate).

We conclude that AssistDent® Artificial Intelligence software significantly improves dentists' ability to detect enamel-only proximal caries, with only a slight increase in false positives, and could be considered as a tool to support minimum intervention and preventive dentistry in general practice.

Introduction

The early detection and treatment of enamel-only proximal caries can preserve tooth structure and prevent the subsequent cycle of treatment and re-treatment that is involved with more invasive treatment. Recent guidelines from the NHS encourage preventive care in dental practices, especially for young children. Patients differ widely on their willingness to pay for preventive therapies[1] but nearly all parents value a healthy dentition for their children and are willing to invest resources to maintain this.[2] However, it is well documented that preventive care in adults is offered less frequently than it should be and therefore if prevention is to be adopted more widely in adults, caries detection must be time efficient and accurate. Only then can the ideal, personalised caries assessment of adult patients and their preventive care be developed.

Radiographic examinations can increase the number of carious lesions that are detected and that would not be detectable by clinical examination alone. Nevertheless, systematic reviews have consistently shown that detection of proximal caries on bitewing radiography has a low sensitivity.[3] A number of studies have reported poor diagnostic sensitivity for radiographic detection of demineralisation by dentists. In a classic study by Mejàre et al.[4], premolar and adjacent teeth surfaces were examined radiographically and visually. The premolar teeth were then extracted for orthodontic reasons. They found that the sensitivity of detection of enamel-only proximal caries was 36.7%. Other studies have found similarly low sensitivity values.[5,6]

The introduction of Artificial Intelligence (AI) methods allows assessment of routine tasks to be conducted more quickly and efficiently. AssistDent® is an AI software product, developed by Manchester Imaging Limited, and uses machine learning algorithms to search for evidence of enamel-only proximal caries on bitewing radiographs.[7] It is an aid to the dentist, assisting their clinical decision-making by provided on-screen prompts for potential locations of enamel-only proximal caries. The final judgement about whether enamel-only proximal caries is present, or not, is a decision for the clinician. The null hypothesis of this research was that there is no difference in the performance of dentists in diagnosing the presence of enamel-only proximal caries on bitewing radiographs with and without the use of AssistDent®.

Methods

A pilot study[8] was conducted with dental students to assist in developing the methodology and provide initial data for a sample size calculation. Final study protocol, participant information sheet and consent forms for the study were approved by the Manchester University Research Ethics Committee (Ref: 2020-9892-15955).

Participants were recruited from two sources: 1) dentists practising as general dental practitioners who in addition provide tutorage for dental students within the University of Manchester Dental School; 2) practising dentists undertaking postgraduate training within the University of Manchester Foundation NHS trust. All signed informed consent forms.

The dentists were randomly divided into control and experimental arms by pairing the participants within the recruitment sources according to the order in which they were enrolled. The first of each pair of participants were randomly assigned to either the control or experimental arm with the second assigned to the other group. This method ensured random assignment while maintaining even arm sizes equally balanced between the recruitment sources as the study progressed.

Both arms examined the same images using the same graphical user interface. In the control group (n=11), the caries prompting function of AssistDent® was disabled in order to measure the ability of the group to detect enamel-only proximal caries without the use of AI software. In the experimental group (n=12), the caries prompting function of AssistDent® was enabled in order to assist the participants.

So that the study would have applicability to general practice, a total of 1,446 bitewing radiographs were collected from a range of different sites (1 teaching hospital site and 9 general dental practitioner (GDP) sites). Separate ethical approval had been received from the Integrated Research Application System (IRAS project ID: 248306, REC reference: 18/NI/0111). A validation set of 103 images were selected by random stratified sampling partitioned over the image acquisition sites and excluded from all machine learning model training and evaluation. A further subset of 24 images from the validation set were selected for the study, again stratified over the acquisition sites but with the criterion that there was at least one enamel-only proximal caries in each image. Images from one of the GDP sites were excluded due to their poor quality. Another GDP site lacked images with one enamel-only proximal caries, therefore two study images from this site had no enamel-only proximal caries. The images were presented to each participant in the same order, grouped according to the acquisition site.

Gold Standard annotation of all classes of proximal caries was obtained from a panel of 5 dento-maxillofacial radiologists and 1 Professor of Restorative Dentistry, each of whom annotated the location and grade of caries on a set of images. These individual expert annotations were consolidated resulting in a gold standard set of 1,972 examples of enamel-only proximal caries for algorithm training and evaluation.

The caries annotations entered by each participant were collected remotely via a web application and analysed to determine whether they were true positives (correct identifications of enamel-only proximal caries) or false positives (annotations not corresponding to the location of the gold standard enamel-only proximal caries). Annotations corresponding to dentine proximal caries were recorded but excluded from this analysis. The **True Positive Rate** or **Sensitivity** of diagnosis is a measure of how well a participant detected the enamel-only proximal caries and was calculated as the sum of true positives divided by the sum of the gold standard caries. **True Negative Rate** or **Specificity** is a measure of how well the participant identified healthy surfaces and did not mark them as carious. The probability of a false detection is quoted in terms of a **False Positive Rate** calculated as the sum of false positive detections divided by the sum of healthy surfaces which is equal to 1-Specificity.

Results

Table 1 presents a per-participant breakdown of the evaluation scores and performance measures together with the aggregate scores and measures for each arm. The data demonstrates that 23 dentists were recruited, 11 in the control arm and 12 in the experimental arm. These were balanced within the arms between the two recruitment sources of general dental practitioners and practising dentists undertaking postgraduate training. All participants analysed all 24 images. The aggregate measures presented at the bottom of the table are calculated as the sum of the quantitative measures across all participants within each arm, together with the aggregate performance measures.

Figure 1 illustrates the True Positive and False Positive Rates of each participant on a scatter plot. The plot illustrates the close grouping of both true positive and false positive rates for the control arm. True positive rate is also tightly grouped for the experimental arm whereas there is a greater range in false positive rates which are generally worse than those for the control arm.

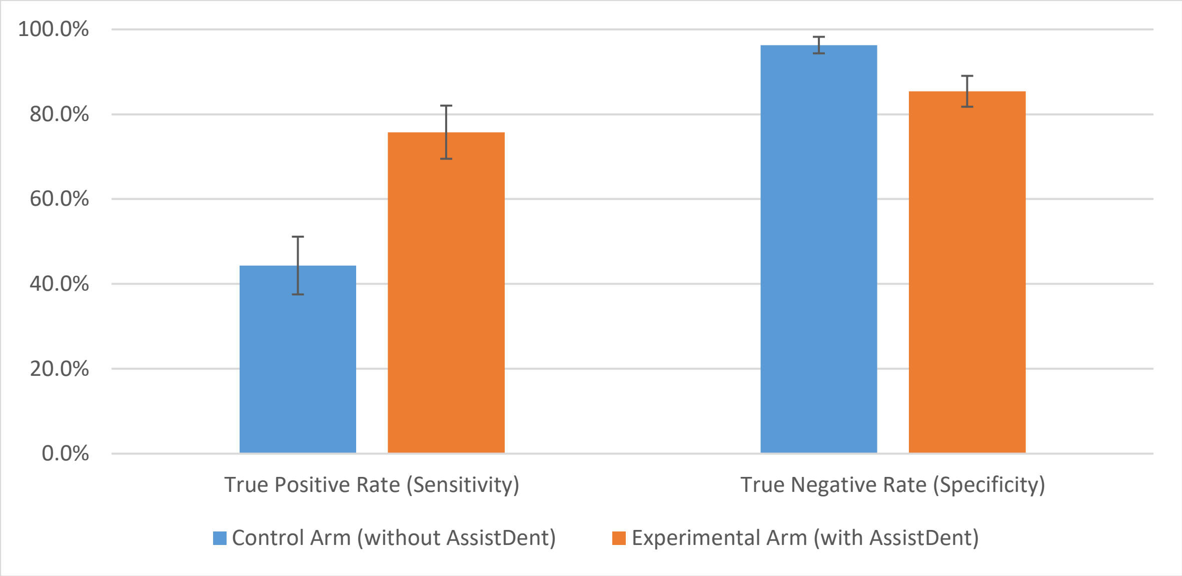
Figure 2 presents the mean true positive and true negative rates over all participants and the 95% confidence intervals, for each arm together with the t-test analysis. The improved mean true positive rate of the experimental arm participants (75.8% with AssistDent®) compared to the control arm (44.3% without AssistDent®) is clearly visible. This is accompanied by a decrease in true negative rate from 96.3% to 85.4%. The table in Figure 2 presents the statistical analysis for the performance measures within each arm together with the result of a student t-test. The t-tests demonstrate that the improved true positive rate and reduced true negative rate of the experimental compared to the control arm were significant with p-values below alpha of 0.01.

Table 2 presents the results of an odds-ratio comparison of true positive and true negative rates for each arm. The ratio between the experimental and control groups with a value greater than one indicates that use of AssistDent® increased the ability to detect enamel-only proximal caries by 71%. Similarly, the ratio of less than 1 for true negative rate indicates that the experimental arm participants were 11% less likely to correctly identify healthy proximal surfaces as non-carious.

Participant	True Positives	False Positives (Healthy marked as Diseased)	True Positive Rate (Sensitivity)	True Negative Rate (Specificity)	False Positive Rate (1-Specificity)
Control Arm (without AssistDent®)					
C1	22	3	0.338	0.988	0.012
C2	29	13	0.446	0.946	0.054
C3	42	28	0.646	0.884	0.116
C4	23	6	0.354	0.975	0.025
C5	22	9	0.338	0.963	0.037
C6	30	6	0.462	0.975	0.025
C7	34	6	0.523	0.975	0.025
C8	27	5	0.415	0.979	0.021
C9	23	4	0.354	0.983	0.017
C10	28	11	0.431	0.954	0.046
C11	37	7	0.569	0.971	0.029
Total	**317**	**98**	**0.443**	**0.963**	**0.037**
Experimental Arm (with AssistDent®)					
E1	52	41	0.800	0.830	0.170
E2	53	44	0.815	0.817	0.183
E3	40	17	0.615	0.929	0.071
E4	55	55	0.846	0.772	0.228
E5	50	24	0.769	0.900	0.100
E6	51	38	0.785	0.842	0.158
E7	53	33	0.815	0.863	0.137
E8	39	20	0.600	0.917	0.083
E9	39	12	0.600	0.950	0.050
E10	49	45	0.754	0.813	0.187
E11	52	45	0.800	0.813	0.187
E12	58	48	0.892	0.801	0.199
Total	**591**	**422**	**0.758**	**0.854**	**0.146**

Table 1 Per-participant breakdown and aggregate performance measures for each arm. Every participant analysed all 24 images meaning that they were all exposed to 65 enamel-only proximal caries and 241 healthy surfaces.

Figure 1 Scatter plot of the True Positive Rate (Sensitivity) versus False Positive Rate (1-Specificity) for each participant coloured according to arm.



	True Positive Rate (Sensitivity)		True Negative Rate (Specificity)	
	Control Arm (without AssistDent®)	Experimental Arm (with AssistDent®)	Control Arm (without AssistDent®)	Experimental Arm (with AssistDent®)
n	11	12	11	12
Mean	44.3%	75.8%	96.3%	85.4%
Standard Deviation	10.1%	9.9%	2.9%	5.7%
95% Confidence Interval	±6.8%	±6.3%	±1.9%	±3.6%
t-test p-value	2.6×10^{-8}		2.6×10^{-6}	

Figure 2 Bar charts showing the mean true positive and true negative rates together with their 95% confidence intervals, for each arm together with statistical analysis of the mean per-participant performance measures for each arm.

	True Positive Rate (Sensitivity) Detection of Enamel-Only Proximal Caries			*True Negative Rate (Specificity) Identification of Healthy Surfaces*		
	Detected	*Undetected*	*Total*	*Detected*	*Undetected*	*Total*
Experimental	*591*	*189*	*780*	*2470*	*422*	*2892*
Control	*317*	*398*	*715*	*2553*	*98*	*2651*
Correctly identified in Experimental	*75.8%*			*85.4%*		
Correctly identified in Control	*44.3%*			*96.3%*		
Relative Risk (95% Confidence Interval)	*1.71 (1.56, 1.87)*			*0.89 (0.87,0.90)*		

Table 2 Odds-Ratio of true positive and true negative rates of the experimental group compared to the control group with 95% confidence intervals.

Discussion

Small areas of proximal surface enamel demineralization are difficult to detect visually on radiographs.[9] In a review of diagnostic studies reported by Keenan and Keenan,[10] they found lower sensitivities for the detection of proximal lesions in clinical than in vitro studies. The mean sensitivity in the clinical studies was 0.24 and for the in vitro studies was 0.43.

Other studies that have measured sensitivity and specificity of caries detection using radiographs have found a large variation. This may be due to variation in the range of caries depth or the use of in vitro studies. In our in vivo study, we mainly used a sample of radiographs from a range of dental practices that were taken as part of the routine examination of patients. The low sensitivity and high specificity of our control group in detecting caries on bitewing radiographs are comparable with the above studies, and with that reported in a systematic review.[11]

Low sensitivity in the detection of enamel-only caries is problematic in the context of minimally-invasive dentistry. A significant increase in sensitivity is required over the values reported in these studies, even at the expense of an increased number of false positive detections. Enamel lesions are treated in a conservative manner with dietary and interdental cleaning advice, fluoride treatments and potentially, resin infiltration. This, in conjunction with cooperation from the patient will prevent operative intervention and the entry of the tooth into a restorative cycle of increasingly larger restorations and eventual extraction. The relatively small decrease in specificity observed from the use of AssistDent® accompanies a much larger increase in sensitivity.

There has been a growing interest in general dental publications and online fora in the use of AI in dentistry. A recent review by Schwedicke et al [12] has described the basics of AI and explored its potential use in diagnostics, treatment planning and conduct in dentistry. Computer image analysis is a particular application of AI in dentistry (and medicine more generally). An earlier, scoping review [13] refers particularly to the application in image diagnostics. There is a relatively small number of studies related to caries detection. For example, Lee et al.[14] investigated the performance of a convolutional neural network on classification of caries in images of individual teeth isolated from periapical images. The study did not focus specifically on enamel-only caries. Srivastra et al.[15] also trained a convolutional neural network for fully automatic detection of caries in bitewing images.

To our knowledge, AssistDent® is the only commercially available AI system for use in the clinical diagnosis of enamel-only proximal caries, acting as a prompting system to support dentists' diagnostic decisions. Schwendicke et al [13] recommend that "…the dental community should appraise [the AI systems] against the rules of evidence-based practice." This study is an example of such an appraisal.

It is important that dentists receive appropriate training with any new diagnostic system. Qudeimat et al.[16] investigated the effect of ICDAS training and found a significant increase in overtreatment recommendations. The reproducibility of any diagnostic system depends on the experience of the clinician and this is especially so where the diagnosis is a visual score rather than AI-based.[17]

How should our profession address the issue of poor sensitivity in detection of early enamel caries? Audit, reflection with peer review and evaluation of past performance are essential parts of dental practice. An audit of caries diagnosis will rely on identifying opportunities for improvement, comparison with an accepted standard of care and implementing change. The increased sensitivity arising from AI-supported detection may provide a useful standard for audit in caries detection. Testing will also reveal clinicians with unacceptable variation in bitewing analysis competency assessments, with direction towards further training. More efficient caries detection may help in identifying patients with higher caries risk, while the associated display of detected caries can provide a basis for encouraging a detailed discussion with the patient of their oral health and the factors affecting it. The intuitive graphical screen display of AssistDent® may be more readily understood by patients in comparison to other algorithm-based tools such as the Cariogram. The latter uses a pie-diagram of important factors to illustrate the probability that future caries may be prevented.[19] However, once the high caries risk individuals are identified, it is important to further investigate all the factors predisposing to caries such as a high sugar intake and infrequent brushing.[19]

Caries assessment by dentists using AssistDent® is compatible with the ICCMS caries management system of the ICDAS Foundation and could be used in conjunction with it. AssistDent® provides feedback on the proximal enamel surfaces using radiographs whereas ICCMS is a mainly visual assessment of the non-proximal surfaces for caries. Both systems aim to maintain tooth structure and encourage preventive care by developing a caries assessment for each patient.

Conclusion

The small increase in false positive diagnoses that arises from the AI software AssistDent® may result in an increase of unnecessary preventative treatment and associated use of limited personal and public resources.

However, the significantly greater increase in sensitivity in detecting enamel-only proximal caries should enable increased application of minimally invasive procedures. More accurate targeting of preventive treatments would thereby avoid the requirement for later restoration, potentially resulting in an overall saving of resources and an improvement in the dentition of patients.

Declaration of interests

HD, JG and TW are employees, of Manchester Imaging Ltd. The Division of Dentistry, University of Manchester, purchased a software licence for AssistDent® from Manchester Imaging Ltd. MA is not an employee of Manchester Imaging Ltd and declares no conflict of interest.

References

1. Vernazza CR, Wildman JR, Steele JG, et al. Factors affecting patient valuations of caries prevention: Using and validating the willingness to pay method. J Dent. 2015; **43**: 981-8.

2. Berendsen J, Bonifacio C, van Gemert-Schriks M, et al. Parents' willingness to invest in their children's oral health. J Public Health Dent. 2018; **78**: 69-77.

3. Ismail AI. Visual and visuo-tactile detection of dental caries. J Dent Res 2004; **83**(Spec No C): C56–66.

4. Mejàre I, Gröndahl HG, Carlstedt K, Grever AC, Ottosson E Accuracy at radiography and probing for the diagnosis of proximal caries. Scand J Dent Res. 1985; **93**: 178-84.

5. Machiulskiene V, Nyvad B, Baelum V. A comparison of clinical and radiographic caries diagnoses in posterior teeth of 12-year-old Lithuanian children. Caries Res 1999; **33**: 340–8.

6. Machiulskiene V, Nyvad B, Baelum V. Comparison of diagnostic yields of clinical and radiographic caries examinations in children of different age. Eur J Paediatr Dent 2004; **5**: 157–62.

7. AssistDent Artificial Intelligence Software. <https://www.assistdent.net>.

8. Devlin H, Ashley M, Williams TG, Purvis B. A Pilot Comparative Study of Dental Students' Ability to Detect Enamel-only Proximal Caries in Bitewing Radiographs With and Without the use of AssistDent® Deep Learning Software medRxiv preprint doi https://doi.org/10.1101/2020.06.15.20131730 ,June 2020

9. Ismail AI, Sohn W, Tellez M, Amaya A, Sen A, Hasson H, et al. The International Caries Detection and Assessment System (ICDAS): an integrated system for measuring dental caries. Community Dent Oral Epidemiol 2007; **35**: 170-8.

10. Keenan JR, Keenan AV. Accuracy of dental radiographs for caries detection. Evid Based Dent. 2016; **17**: 43.

11. Schwendicke F, Tzschoppe M, Paris S. Radiographic caries detection: a systematic review and meta-analysis. J Dent 2015; **43**: 924–33.

12. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: Chances and Challenges. J Dental Research 2020; **99**: 769-774.

13. Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: A scoping Review. J Dentistry 2019; 102 103226. Available online at doi: 10.1016/j.jdent.2019.103226. Epub 2019 Nov 5. PMID: 31704386.

14. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent. 2018; **77**: 106-111.

15. Srivastava MM, Kumar P, Pradham L, Varadarajam S. Detection of tooth caries in bitewing images using deep learning. ArXiv 1711.07312v1 [cs.CV] 20 Nov 2017.

16. Qudeimat MA, Altarakemah Y, Alomari Q, Alshawaf N, Honkala E. The impact of ICDAS on occlusal caries treatment recommendations for high caries risk patients: an in vitro study. BMC Oral Health. 2019; **19**: 41.

17. Qudeimat MA, Alomari QD, Altarakemah Y, Alshawaf N, Honkala EJ. Variables affecting the inter- and intra-examiner reliability of ICDAS for occlusal caries diagnosis in permanent molars. J Public Health Dent. 2016; **76**: 9-16.

18. Petersson GH. Assessing caries risk--using the Cariogram model. Swed Dent J Suppl. 2003; **158**: 1-65.

19. Llena C, Calabuig E. Risk factors associated with new caries lesions in permanent first molars in children: a 5-year historical cohort follow-up study. Clin Oral Investig. 2018; **22**: 1579-1586.