University of Dundee

Transillumination and optical coherence tomography for the detection and diagnosis of enamel caries

Macey, Richard; Walsh, Tanya; Riley, Philip; Hogan, Richard; Glenny, Anne-Marie; Worthington, Helen V.

Published in:

The Cochrane database of systematic reviews

DOI:

[10.1002/14651858.CD013855](https://doi.org/10.1002/14651858.CD013855)

Publication date:

2021

Document Version

Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Macey, R., Walsh, T., Riley, P., Hogan, R., Glenny, A-M., Worthington, H. V., Clarkson, J. E., & Ricketts, D. (2021). Transillumination and optical coherence tomography for the detection and diagnosis of enamel caries. *The Cochrane database of systematic reviews*, 2021(1), [CD013855]. <https://doi.org/10.1002/14651858.CD013855>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Cochrane
Library

Cochrane Database of Systematic Reviews

Transillumination and optical coherence tomography for the detection and diagnosis of enamel caries (Review)

Macey R, Walsh T, Riley P, Hogan R, Glenny AM, Worthington HV, Clarkson JE, Ricketts D

Macey R, Walsh T, Riley P, Hogan R, Glenny A-M, Worthington HV, Clarkson JE, Ricketts D.
Transillumination and optical coherence tomography for the detection and diagnosis of enamel caries.
Cochrane Database of Systematic Reviews 2021, Issue 1. Art. No.: CD013855.
DOI: [10.1002/14651858.CD013855](https://doi.org/10.1002/14651858.CD013855).

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS	4
BACKGROUND	6
Figure 1.	8
OBJECTIVES	9
METHODS	9
RESULTS	13
Figure 2.	14
Figure 3.	16
Figure 4.	17
Figure 5.	19
Figure 6.	20
Figure 7.	21
Figure 8.	22
Figure 9.	23
Figure 10.	24
Figure 11.	25
Figure 12.	26
Figure 13.	27
Figure 14.	28
Figure 15.	29
Figure 16.	30
Figure 17.	31
Figure 18.	32
DISCUSSION	33
Figure 19.	35
AUTHORS' CONCLUSIONS	36
ACKNOWLEDGEMENTS	37
REFERENCES	38
CHARACTERISTICS OF STUDIES	45
DATA	101
Test 1. ALL	102
Test 2. NIR	102
Test 3. OCT	102
Test 4. FOTI/DIFOTI	103
Test 5. DIFOTI	103
ADDITIONAL TABLES	103
APPENDICES	108
HISTORY	109
CONTRIBUTIONS OF AUTHORS	109
DECLARATIONS OF INTEREST	109
SOURCES OF SUPPORT	110

[Diagnostic Test Accuracy Review]

Transillumination and optical coherence tomography for the detection and diagnosis of enamel caries

Richard Macey¹, Tanya Walsh¹, Philip Riley², Richard Hogan³, Anne-Marie Glenny¹, Helen V Worthington², Janet E Clarkson⁴, David Ricketts⁵

¹Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. ²Cochrane Oral Health, Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. ³Dental Health Unit, Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. ⁴Division of Oral Health Sciences, Dundee Dental School, University of Dundee, Dundee, UK. ⁵Dundee Dental School, University of Dundee, Dundee, UK

Contact address: Tanya Walsh, tanya.walsh@manchester.ac.uk.

Editorial group: Cochrane Oral Health Group.

Publication status and date: New, published in Issue 1, 2021.

Citation: Macey R, Walsh T, Riley P, Hogan R, Glenny A-M, Worthington HV, Clarkson JE, Ricketts D. Transillumination and optical coherence tomography for the detection and diagnosis of enamel caries. *Cochrane Database of Systematic Reviews* 2021, Issue 1. Art. No.: CD013855. DOI: [10.1002/14651858.CD013855](https://doi.org/10.1002/14651858.CD013855).

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Caries is one of the most prevalent and preventable conditions worldwide. If identified early enough then non-invasive techniques can be applied, and therefore this review focusses on early caries involving the enamel surface of the tooth. The cornerstone of caries detection and diagnosis is a visual and tactile dental examination, although alternative approaches are available. These include illumination-based devices that could potentially support the dental examination. There are three categories of illumination devices that exploit various methods of application and interpretation, each primarily defined by different wavelengths, optical coherence tomography (OCT), near-infrared (NIR), and fibre-optic technology, which incorporates more recently developed digital fibre optics (FOTI/DIFOTI).

Objectives

To estimate the diagnostic test accuracy of different illumination tests for the detection and diagnosis of enamel caries in children or adults. We also planned to explore the following potential sources of heterogeneity: in vitro or in vivo studies with different reference standards; tooth surface (occlusal, proximal, smooth surface, or adjacent to a restoration); single or multiple sites of assessment on a tooth surface; and the prevalence of caries into dentine.

Search methods

Cochrane Oral Health's Information Specialist undertook a search of the following databases: MEDLINE Ovid (1946 to 15 February 2019); Embase Ovid (1980 to 15 February 2019); US National Institutes of Health Ongoing Trials Register (ClinicalTrials.gov, to 15 February 2019); and the World Health Organization International Clinical Trials Registry Platform (to 15 February 2019). We studied reference lists as well as published systematic review articles.

Selection criteria

We included diagnostic accuracy study designs that compared the use of illumination-based devices with a reference standard (histology, enhanced visual examination with or without radiographs, or operative excavation). These included prospective studies that evaluated the diagnostic accuracy of a single index test and studies that directly compared two or more index tests. Both in vitro and in vivo studies of primary and permanent teeth were eligible for inclusion. We excluded studies that explicitly recruited participants with caries into dentine

or frank cavitation. We also excluded studies that artificially created carious lesions and those that used an index test during the excavation of dental caries to ascertain the optimum depth of excavation.

Data collection and analysis

Two review authors extracted data independently and in duplicate using a standardised data extraction form and quality assessment based on QUADAS-2 specific to the clinical context. Estimates of diagnostic accuracy were determined using the bivariate hierarchical method to produce summary points of sensitivity and specificity with 95% confidence regions. The comparative accuracy of different illumination devices was conducted based on indirect and direct comparisons between methods. Potential sources of heterogeneity were pre-specified and explored visually and more formally through meta-regression.

Main results

We included 24 datasets from 23 studies that evaluated 16,702 tooth surfaces. NIR was evaluated in 6 datasets (673 tooth surfaces), OCT in 10 datasets (1171 tooth surfaces), and FOTI/DIFOTI in 8 datasets (14,858 tooth surfaces). The participant selection domain had the largest number of studies judged at high risk of bias (16 studies). Conversely, for the index test, reference standard, and flow and timing domains the majority of studies were judged to be at low risk of bias (16, 12, and 16 studies respectively). Concerns regarding the applicability of the evidence were judged as high or unclear for all domains. Notably, 14 studies were judged to be of high concern for participant selection, due to selective participant recruitment, a lack of independent examiners, and the use of an in vitro study design. The summary estimate across all the included illumination devices was sensitivity 0.75 (95% confidence interval (CI) 0.62 to 0.85) and specificity 0.87 (95% CI 0.82 to 0.92), with a diagnostic odds ratio of 21.52 (95% CI 10.89 to 42.48). In a cohort of 1000 tooth surfaces with a prevalence of enamel caries of 57%, this would result in 142 tooth surfaces being classified as disease free when enamel caries was truly present (false negatives), and 56 tooth surfaces being classified as diseased in the absence of enamel caries (false positives). A formal comparison of the accuracy according to device type indicated a difference in sensitivity and/or specificity ($\text{Chi}^2(4) = 34.17$, $P < 0.01$). Further analysis indicated a difference in the sensitivity of the different devices ($\text{Chi}^2(2) = 31.24$, $P < 0.01$) with a higher sensitivity of 0.94 (95% CI 0.88 to 0.97) for OCT compared to NIR 0.58 (95% CI 0.46 to 0.68) and FOTI/DIFOTI 0.47 (95% CI 0.35 to 0.59), but no meaningful difference in specificity ($\text{Chi}^2(2) = 3.47$, $P = 0.18$).

In light of these results, we planned to formally assess potential sources of heterogeneity according to device type, but due to the limited number of studies for each device type we were unable to do so. For interpretation, we presented the coupled forest plots for each device type according to the potential source of heterogeneity.

We rated the certainty of the evidence as low and downgraded two levels in total due to avoidable and unavoidable study limitations in the design and conduct of studies, indirectness arising from the in vitro studies, and imprecision of the estimates.

Authors' conclusions

Of the devices evaluated, OCT appears to show the most potential, with superior sensitivity to NIR and fibre-optic devices. Its benefit lies as an add-on tool to support the conventional oral examination to confirm borderline cases in cases of clinical uncertainty. OCT is not currently available to the general dental practitioner, and so further research and development are necessary. FOTI and NIR are more readily available and easy to use; however, they show limitations in their ability to detect enamel caries but may be considered successful in the identification of sound teeth.

Future studies should strive to avoid research waste by ensuring that recruitment is conducted in such a way as to minimise selection bias and that studies are clearly and comprehensively reported. In terms of applicability, any future studies should be undertaken in a clinical setting that is reflective of the complexities encountered in caries assessment within the oral cavity.

PLAIN LANGUAGE SUMMARY

Light-based tests for the detection and diagnosis of early tooth decay

Why is it important to improve dental caries (tooth decay) detection?

Dentists often aim to identify tooth decay that has already advanced to a level which needs a filling. If dentists were able to find tooth decay when it has only affected the outer layer of the tooth (enamel) then it is possible to stop the decay from spreading any further and prevent the need for fillings. It is also important to avoid a false-positive result, when treatment may be provided when caries is absent.

What is the aim of this review?

This Cochrane Review aimed to find out how accurate different forms of light-based tests are for detecting early tooth decay in patients who regularly visit their dentist. Researchers in Cochrane included 23 studies published between 1988 and 2019 to answer this question.

What was studied in the review?

We included three different types of light-based devices in this review: optical coherence tomography (OCT), near-infrared (NIR), and fibre-optic (FOTI/DIFOTI) technology. All devices rely on shining different types of light on the tooth and can improve the dentist's ability to identify tooth decay.

Transillumination and optical coherence tomography for the detection and diagnosis of enamel caries (Review)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

What are the main results of the review?

The review included 23 studies with a total of 16,702 tooth surfaces. The results of these studies indicate that if the illumination devices were used by a dentist for a routine dental examination of 1000 tooth surfaces, of which 570 (57%) have early tooth decay:

- an estimated 484 would be found to have tooth decay using one of the illumination detection methods, and of these 56 (12%) would not have tooth decay (false positive - incorrect diagnosis);
- of the 516 tooth surfaces in which a device indicated that tooth decay is not present, 142 (28%) tooth surfaces will truly have early tooth decay (false negative - incorrect diagnosis).

Please see oralhealth.cochrane.org/transillumination-and-optical-coherence-tomography-detection-and-diagnosis-enamel-caries-results.

In this example illumination devices produce a relatively high proportion of false-negative results, whereby patients do not receive treatment for early tooth decay, for example, high fluoride toothpaste or oral health advice and guidance from the dentist, as they should. Of the data collected from three types of illumination devices, it seems that the OCT device is more sensitive (produces fewer false-negative results) than NIR or fibre-optic technology.

How reliable are the results of the studies in this review?

We only included studies that assessed healthy teeth or those that were thought to have early tooth decay, as teeth with deep tooth decay would be easier to identify. There were some shortcomings in how the studies were conducted, and this may have resulted in the illumination devices appearing more accurate than they really are, increasing the number of correct classifications (green rectangles in the diagram). Many studies evaluated the performance of the devices on extracted teeth, which is very different from when the devices are used inside a person's mouth, where it is difficult to see clearly and where teeth may be stained or have a covering of plaque.

Who do the results of this review apply to?

Studies included in the review were carried out in the United States, Europe, Japan, Brazil, China, Malaysia, and Australia. Others were completed in dental hospitals, general dental practices, or schools.

What are the implications of this review?

Optical coherence tomography (OCT) shows potential as a device to detect early/enamel caries but more high-quality research and development are required as OCT is not currently available to general dental practitioners. The analysis suggests that OCT is superior to NIR and fibre-optic technologies.

How up-to-date is this review?

The review authors searched for and used studies published up to 15 February 2019.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings table

Question	What is the diagnostic accuracy of transillumination-based index tests for the detection and diagnosis of early dental caries?			
Population	Children or adults who are presenting asymptotically or are suspected of having enamel caries (clinical studies); extracted teeth of children or adults (in vitro studies). Studies which intentionally included dentine and frank cavitations were excluded			
Index test	Transillumination-based devices - including near-infrared (NIR), optical coherence tomography (OCT), and fibre-optic transillumination (FOTI)/digital fibre-optic transillumination (DIFOTI), suitable for use as an adjunct to a conventional clinical oral examination. The index tests produced an enhanced view of the tooth and were interpreted by a trained examiner			
Comparator test	Comparisons were made between transillumination devices. A separate review in this series investigates comparisons between transillumination-based tests and enhanced visual, radiograph, and fluorescence tests			
Target condition	Dental caries, at the threshold of caries in enamel			
Reference standard	Histology, enhanced visual examination with or without radiographs			
Action	If dental caries can be detected at an early stage then remedial action can be taken to arrest or reverse the decay and potentially prevent restorations			
Diagnostic stage	Aimed at the general dental practitioner assessing regularly attending patients for early stage caries			
Quantity of evidence	23 studies providing data for meta-analysis (24 datasets, 16,702 teeth, 2499 tooth surfaces with total caries at enamel threshold or greater (15% prevalence))			
Findings	All studies (24 datasets)	NIR (6 datasets, 673 tooth surfaces, 56% prevalence)	OCT (10 datasets, 1171 tooth surfaces, 52% prevalence)	FOTI/DIFOTI (8 datasets, 14,858 tooth surfaces, 10% prevalence)
Sensitivity (95% CI)^a	0.75 (0.62 to 0.85)	0.58 (0.46 to 0.68)	0.94 (0.88 to 0.97)	0.47 (0.35 to 0.59)
Specificity (95% CI)^a	0.87 (0.82 to 0.92)	0.86 (0.80 to 0.91)	0.83 (0.68 to 0.91)	0.92 (0.86 to 0.96)
DOR (95% CI)	21.52 (10.89 to 42.48)	8.65 (3.92 to 19.06)	72.07 (23.87 to 217.66)	10.75 (4.49 to 25.72)
Effect per 1000 tooth surfaces assessed	Numbers applied to a hypothetical cohort of 1000 tooth surfaces (95% CI)		Test accuracy Certainty of the evidence	
Outcome	Pre-test probability 28% ^b	Pre-test probability 57% ^b		
True positives (patients with early enamel caries)	210 (174 to 238)	428 (353 to 484)	⊕⊕⊕○ LOW	

False negatives (patients incorrectly classified as not having early enamel caries)	70 (42 to 106)	142 (86 to 271)
True negatives (patients without early enamel caries)	626 (540 to 662)	374 (353 to 396)
False positives (patients incorrectly classified as having early enamel caries)	94 (58 to 180)	56 (34 to 77)

Limitations - factors that may decrease the certainty of the evidence	
Risk of bias	<p>Of the 23 studies: the participant selection domain had the largest number of studies judged at high risk of bias (16 studies). Conversely, for the index test, reference standard, and flow and timing domains the majority of studies were judged to be at low risk of bias (16, 12, and 16 studies respectively)</p> <p>For the index test domain, 2 studies suggested that the index test may have been influenced by the reference standard, a further 3 lacked independent examiners, and 1 was unclear on whether the threshold was predetermined</p> <p>The reference standard was deemed to have correctly classified the target condition in 12 studies and was interpreted without knowledge of the index test in 12 studies</p> <p>There was no concern regarding the interval between the index test and the reference standards in 21 studies, the same reference standard was used for all tooth surfaces in 23 studies, and all tooth surfaces were reported in the analysis in 12 studies</p>
Applicability of evidence to the review question	<p>High concern was observed for patient selection where extracted teeth were used (14 studies), where the index test applied a device that is not currently available to a general dental practitioner (4 studies), and where the reference standard was a radiograph (4 studies)</p>
Certainty of the evidence	<p>We rated the certainty of the evidence as low and downgraded 2 levels in total due to avoidable and unavoidable study limitations in the design and conduct of studies, indirectness arising from the in vitro studies, and imprecision of the estimates</p>

^aSummary estimates of sensitivity and specificity were reported for all included studies but the 3 groups of devices vary in their design and use so it was necessary to present the results of the subgroups within the 'Summary of findings' table.

^bPrevalence values of 28% and 57% were used to calculate the natural frequencies based on the summary estimates. The prevalence of all assessed surfaces included in this review was 15%, however the overall prevalence was skewed downwards by 3 large studies in the FOTI/DIFOTI group where the prevalence was comparatively low at 3%, 16%, and 19%. The pre-test probabilities of 28% and 57% were used to facilitate comparisons of results with other reviews in this series, and also serve as a more representative prevalence for the 18 NIR and OCT studies included in this review. Based on consultation with clinical colleagues, the lower prevalence value of 28% addresses concerns regarding the representativeness of the overall prevalence value in this review. The 28% value is taken from the level of cavitated teeth reported in the UK Adult Dental Health Survey (Steele 2011). The higher prevalence value is taken from the prevalence of enamel caries in the fluorescence review in this series (Macey 2020).

CI: confidence interval; **DOR:** diagnostic odds ratio.

BACKGROUND

Cochrane Oral Health (COH) has undertaken a number of Cochrane Reviews of diagnostic test accuracy (DTA) on the detection and diagnosis of dental caries. The suite of systematic reviews forms part of a UK National Institute for Health Research (NIHR) Cochrane Programme Grant Scheme and involved collaboration with the Complex Reviews Support Unit. The reviews follow standard Cochrane DTA methodology and will be differentiated according to the index test under evaluation. A generic protocol served as the basis for the suite of systematic reviews (Macey 2018).

Caries is an entire disease process, which can be stabilised and sometimes reversed if diagnosed and treated early on in the disease process (Fejerskov 2015; Pitts 2009). In some Scandinavian countries preventative programmes are in place which nearly eradicate primary caries, but this is continuous day-to-day work and has not been replicated in other populations (Pitts 2017). Despite this the 2015 Global Burden of Disease study identified dental caries as the most prevalent, preventable condition worldwide (Feigin 2016; Kassebaum 2015), affecting 60% to 90% of children and the majority of adults of the world's population (Petersen 2005). Furthermore, despite a reduction in caries in some industrialised countries, the global incidence of untreated caries was reported to be 2.4 billion in 2010 (Feigin 2016; Kassebaum 2015; World Health Organization 2017) and continues to increase year on year. In the UK, the primary reason for childhood (aged 5 years to 9 years) hospital admissions is for the extraction of teeth (Public Health England 2014). Longitudinal studies have shown that those who experience caries early in childhood will have an increased risk of severe caries in later life, the trajectory of disease will be of increased severity (Broadbent 2008; Hall-Scullin 2017).

Untreated caries can lead to episodes of severe pain and infection, which is often treated with antibiotics. Dental anxiety, resulting from the failure to treat caries and the subsequent need for more invasive management, can adversely affect a person's future willingness to visit their dentist, leading to a downward spiral of oral disease (Milsom 2003; Thomson 2000). If left to progress, treatment options are limited to restoration or extraction, requiring repeated visits to a dental surgery or even to a hospital (Featherstone 2004; Fejerskov 2015; Kidd 2004).

The cost of treating caries is high. In the UK alone, the National Health Service (NHS) spends around GBP half a billion every year in treating the disease. Hidden costs also exist, and the related productivity losses are high, estimated at USD 27 billion globally in 2010 (Listl 2015).

Caries detection and diagnosis will usually be undertaken at a routine dental examination, by a general dental practitioner, in patients who are presenting asymptotically. However, caries detection can additionally be employed in secondary care settings, school, or community screening projects and epidemiology or research studies (Braga 2009; Jones 2017). The traditional method of detecting dental caries in clinical practice is a visual-tactile examination often with supporting radiographic investigations. This combination of methods is believed to be successful at detecting caries that has progressed into dentine and reached a threshold where restoration is necessary (Kidd 2004). The detection of caries earlier in the disease continuum could lead to stabilisation of disease or even possible remineralisation of the tooth surface, thus preventing the patient from entering a lifelong cycle of

restoration (Pitts 2017). However, early caries is difficult to detect visually, and the use of radiographs provides limited ability to detect small changes in dental enamel (Ismail 2007).

Detection and diagnosis at the initial (non-cavitated) and moderate (enamel) levels of caries is fundamental in achieving the promotion of oral health and prevention of oral disease (Fejerskov 2015; Ismail 2013). The prevalence of this early caries state is not often reported in dental epidemiology, most reports preferring to focus on cavitated/dental lesions which may be easier to detect. An example of this is the most recent UK Adult Dental Health survey which reported 31% of the sample having untreated caries into dentine (Steele 2011; White 2012) and a US study reporting levels of cavities at 15.3% in 12- to 19-year olds (Dye 2015). However, one UK survey of children identified "clinical decay experience" which incorporates any enamel breakdown and all other form of caries and reported a prevalence of 63% in 15-year olds (Vernazza 2016).

A wide variety of treatment options are available under NHS care at these different thresholds of disease, these include:

- non-operative preventive strategies such as improved oral hygiene, reduced sugar diet, and application of topical fluoride;
- minimally invasive treatments such as sealing the affected surface of the tooth, or 'infiltrating' the demineralised tissue with resins;
- operative interventions - step-wise caries removal and restoration for extensive lesions.

With advances in technology over the last two decades, alternative methods of detection have become available, such as digital radiography and the development of fluorescence, transillumination, optical coherence tomography, and electrical conductance devices. These could potentially aid the detection and diagnosis of caries at an early stage of decay. This would afford the patient the opportunity of a less invasive treatment with less destruction of tooth tissue and potentially result in a reduced cost of care to the patient and to healthcare services.

Target condition being diagnosed

Caries is an entire disease process, which can be arrested and sometimes reversed if diagnosed early enough (Fejerskov 2015; Pitts 2009). The term dental caries is used to describe the mechanism which can ultimately lead to the breakdown of the tooth surface which results from an imbalance in the activity within the biofilm (or dental plaque) on the surface of the tooth within the oral cavity (Kidd 2016). This imbalance is especially related to pH levels which fall due to the production of acid by bacteria in the biofilm when exposed to dietary sugars. Disease progression can be moderated by the influx of fluoride through toothpaste and other available fluoride sources. However, the levels of sugar consumption observed in many populations will often outweigh the benefits of fluoride (Hse 2015). Ultimately, carious lesions may develop and destroy the structure of the tooth.

The most common surfaces for caries to manifest are on the biting (occlusal) surface or the tooth surface which faces an adjacent tooth (approximal surfaces); although smooth surfaces adjacent to the tongue, cheeks, and lips can be affected. The severity of disease is defined by the depth of demineralisation of the tooth's structure and whether the lesion is active or arrested. Caries presenting at levels into tooth enamel have potential to be stabilised or

even reversed, whereas the progression of carious lesions into the dentine and pulp of the tooth will often require restoration (Bakhshandeh 2018; Kidd 2004).

Assessment of disease severity traditionally used in epidemiological and research studies has employed some variant of the DMFT (decayed, missing, and filled teeth) scale. Within the D (decayed) component there are four clinically detectable thresholds applied as indicators for diagnosis and treatment planning, often labelled as D₁, D₂, D₃, and D₄ (Anaise 1984) (Additional Table 1). Typically the D₃ threshold has been used to determine the presence of caries (Pitts 1988; Shoaib 2009). These four categories have formed the basis for expanded visual-tactile indices such as the International Caries Detection and Assessment System (ICDAS) (Ekstrand 2007; Ismail 2007).

Treatment of caries

There are many varied treatment options available to the dental clinician, dependent on the thresholds of observed disease. Initial caries can be treated without surgical intervention using preventive and remineralising approaches such as plaque control, dietary advice, and application of fluoride (Kidd 2016). Minimally invasive treatments for initial caries are available, such as sealing the affected surface of the tooth, or 'infiltrating' the porous demineralised tissue with resins. High-risk patients with severe caries may require step-wise caries removal and restoration of extensive lesions.

A caries management pathway, informed by diagnostic information, can be beneficial in guiding the clinician towards prevention or a treatment plan. One recently developed care pathway is the International Caries Classification and Management System (ICCMS) (Ismail 2015). The system presents three forms of management in the care pathway:

- when the dentition is sound the clinician proceeds with preventative strategies to prevent sound surfaces from developing caries;
- non-invasive treatment of the lesion to arrest the decay process and encourage remineralisation, preventing initial lesions from progressing to cavitated decay; and
- management of more severe caries through excavation with tooth preserving (minimally invasive) operative procedures and restoration.

At the core of this care pathway is the ability to detect early caries accurately and optimise the preventative strategies. The detection and diagnosis of early caries remain challenging, and the likelihood of undiagnosed early disease is high (Ekstrand 1997). In such instances, the opportunity for preventing initial lesions from progressing to cavitated decay, or even reversing the disease process, is missed, and disease progresses to cavitated decay where restoration is required (Ekstrand 1998).

Index test(s)

The cornerstone of caries detection is a visual and tactile dental examination, and the ability of clinicians to accurately detect disease in this way has been researched for over half a century (Backer Dirks 1951). A range of additional tests exist which may be suitable at different stages of the care pathway, in particular

focusing on the detection and diagnosis of disease (Bloemendal 2004; Fyffe 2000).

This Cochrane Review investigates illumination-based devices which emit varying wavelengths of light which are transmitted through the tissue of the tooth with results being observed and recorded in a number of ways.

Studies that investigated a standard clinical oral examination with an adjunct of illumination were included if the diagnostic information relating to illumination could be isolated from the other test.

We included four categories of illumination index tests, although it was the intention that other newly emerging devices could be included if they were identified within the searches of the literature.

- Fibre-optic transillumination (FOTI): high intensity white light directed tangentially to the tooth surface of interest with the operating light turned off. Transmitted light through the tooth tissues undergoes light scattering within the pores of a carious lesion more so than sound tooth tissue and the lesion appears as a dark shadow. The results are interpreted qualitatively by the examiner, normally against a classification system which will determine sound, enamel, and dentine lesions (Hogan 2019).
- Digital fibre-optic transillumination (DIFOTI): uses the same approach as FOTI but adds a sensor which allows for an image to be captured with a subjective interpretation still required by a clinician/examiner (Hogan 2019).
- Near-infrared (NIR) transillumination: uses the same approach as DIFOTI but by applying near-infrared light (wavelength of around 780 nm) there is potential for deeper penetration of the tooth structure. The images still require interpretation by an examiner to determine the level of caries (Kühnisch 2019). Commercially available devices include DIAGNOcam or CariVu (KaVo, Biberach, Germany) and VistaCam (Durr Dental, Bietigheim-Bissingen, Germany), and has recently been included in two intraoral scanning technologies (iTero Element 5D (Align Technologies, San Jose USA) and TRIOS 4 (3Shape, Copenhagen, Denmark)).
- Optical coherence tomography (OCT): creates cross-sectional images of the tooth structure through reflection and back scattering of light to measure the optical reflectivity with a depth recording (Fried 2019). Although not currently available to the general dental practitioner, it is anticipated that this emerging diagnostic test will become available in the near future and a synthesis of the available evidence would be of value.

Where combinations of index tests were used (such as radiograph plus illumination), they would be included in this review as a subgroup.

Clinical pathway

The process proceeding from a dental patient attending for a routine examination and a caries assessment being undertaken potentially has four intertwined stages: screening, detection, diagnosis, and treatment planning. If the presenting patient is at some risk of disease but seemingly asymptomatic then this can be considered as a screening exercise (Wilson 1968) to detect initial caries in individuals who do not yet have symptoms. Since caries is a dynamic process the pure detection of the disease at a single time point is not sufficient to inform the future

care of the patient, and additionally the depth and severity of demineralisation, allied to a decision on the caries activity levels, must be combined to reach a diagnosis (Ismail 2004; Nyvad 1997). This diagnosis then feeds into a caries management pathway once the patient's history, personal oral hygiene, and risk factors have been considered. A comprehensive methodology has been developed titled the International Caries Classification and Management System (ICCMS™) which aims to address the need for guidance when diagnosing caries and then following

a decision-making process to use preventative measures and minimise invasive treatment (Ismail 2015). The ICCMS has been developed further into a CariesCare practice guide for dentists to use in primary care to help them provide high-quality care and outcomes for their patients (Martignon 2019).

Figure 1 presents the key elements of the ICCMS process and these reviews could inform the process at 'Keystone 3' where diagnosis is an indefinable component.

Figure 1. keystones of the International Caries Classification and Management System (ICCMS™).

Copyright© 2018 Ismail AI, Pitts NB, Tellez M. The International Caries Classification and Management System (ICCMS™) an example of a caries management pathway. *BMC Oral Health* 2015;15(Suppl 1):S9. Reproduced with permission.

Role of index test(s)

Given that a visual-tactile inspection is the cornerstone of a clinical examination it is unlikely that any of the index tests under evaluation would be used as a complete replacement for the detection and diagnosis of initial decay. Supplementing the visual-tactile examination with an index test could aid in the detection of initial decay. The index tests could also have a triage role in assisting the general dentist to more accurately assess signs of uncertain clinical significance. The information from caries detection (including assessment of severity of disease) will be an integral part of diagnosis, which additionally incorporates patient risk factors and treatment planning protocols.

Alternative test(s)

Alternatives to transillumination and optical coherence tomography include.

- Fluorescence: the breakdown of enamel alters the characteristics of its structure, when exposed to light-inducing fluorescence diseased teeth respond differently to sound teeth. There is potential for mineral loss to be quantified and used to aid the diagnostic decision and treatment pathway (Angmar-Månsson 2001; Matos 2011). Fluorescence is typically divided into laser fluorescence and light fluorescence (i.e. DIAGNOdent type devices and quantitative light-induced fluorescence (QLF) type devices).
- Comprehensive visual or visual-tactile examination with a detailed classification system: identifying caries according to visual appearance, aided by a dental mirror and sometimes a probe, on clean and dry teeth.
- Radiography: bitewing radiology is the most commonly used method. Others techniques include: subtraction radiography which produces a semi-automated method for monitoring progression of lesions (Ellwood 1997; Wenzel 2004) and cone beam computed tomography (CBCT) which provides a three-

dimensional image which appears to offer great potential for diagnosis with increased levels of radiation (Horner 2009).

- Electrical conductance: the demineralisation of the tooth is reported to affect the tooth's electrical conductance. This is measured by placing a probe on the tooth which measures any potentially higher conductivity which occurs due to carious lesions being filled with saliva (Tam 2001).

The relative advantages and disadvantages of each are detailed in Additional Table 2.

Rationale

Despite technological advancement, the current method of caries detection is based upon information from visual-tactile clinical examination with or without radiographs. There have been a number of systematic reviews conducted in this area, Bader 2002 completed an extensive review of in vitro studies investigating visual, radiographic, fibre-optic, electrical conductance, and fluorescence in primary and permanent dentition. Four FOTI studies were included in the review, although three of these assessed caries into dentine. More recently, a review focused on non-cavitated carious lesions and included three studies that investigated FOTI but found the results and quality of the evidence to be poor (Gomez 2013). These reviews predate the development of meta-analysis methods for DTA reviews recommended in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (Deeks 2013). In this Cochrane Review we have built upon existing research in this area to account for the newer technologies not available at the time of the previous reviews by: expanding the search strategy to capture all relevant evidence, for tests developed and used globally; applying appropriate hierarchical meta-analysis (Dinnes 2016); and assessing the body of evidence using GRADE (Schünemann 2020; Schünemann 2020a) to facilitate the production of 'Summary of findings' tables.

OBJECTIVES

To undertake a Cochrane Review of diagnostic test accuracy (DTA) to establish the accuracy of different transillumination and optical coherence tomography (OCT) for the detection and diagnosis of enamel caries.

The specific research question of this review was:

- what is the diagnostic test accuracy of transillumination and OCT for the detection or diagnosis of enamel caries in different populations (children: primary/mixed dentition, adolescents: immature permanent dentition, or adults: mature permanent dentition), and when tested against different reference standards.

Where sufficient studies were available, we planned to evaluate the comparative accuracy of different index tests such as fibre-optic transillumination (FOTI), OCT, and near-infrared (NIR) transillumination.

Secondary objectives

Areas of potential heterogeneity would be investigated.

- In vitro or in vivo studies which may affect the applicability of the results as the laboratory-based studies would not incur the difficulties of examining dentition within the oral cavity. In

vitro versus in vivo studies will also be associated with choice of reference standard - in vitro studies typically use a histological reference standard whereas in vivo studies typically employ excavation or enhanced visual examination.

- Tooth surface (occlusal, proximal, smooth surface or adjacent to a restoration).
- Multiple or single sites of assessment on a tooth surface.
- Prevalence of caries into dentine, where a higher prevalence may have a significant effect on the ability to detect and diagnose caries.
- Recruited population: children: primary/mixed dentition, adolescents: immature permanent dentition, or adults: mature permanent dentition.

METHODS

Criteria for considering studies for this review

Types of studies

We included diagnostic accuracy study designs that were:

- studies with a single set of inclusion criteria that compared a diagnostic test with a reference standard. We included prospective studies that evaluated the diagnostic accuracy of single index tests, and studies that directly compare two or more index tests;
- randomised controlled trials (RCTs) of the diagnostic test accuracy of one or more index tests in comparison, or versus a no test option;
- 'case-control' type accuracy studies where different sets of criteria were used to recruit those with or without the target condition, although prone to bias some innovative tests may be identifiable through this design only and this may provide an opportunity to report them, these studies would not have been included in the primary analysis;
- studies reporting at both the patient and tooth or tooth surface level were included, however only those reporting at the tooth surface level were included in the primary analysis. In vitro and in vivo studies were considered.

In vitro studies are those in which teeth have been extracted prior to the start of the study, commonly for orthodontic purposes, and where caries status was still to be determined. The index test was then performed, albeit in a scenario which is not representative of the typical clinical setting, this would often be followed by a reference standard of histology. In vivo studies recruit apparently healthy participants and conduct index tests and reference standards with the teeth in the oral cavity, without extraction of the teeth and therefore histology would not usually be undertaken but could be on teeth indicated for extraction or primary teeth which were close to exfoliation.

Studies were ineligible for inclusion where:

- artificially created carious lesions were used in the testing procedure;
- an index test was used during the excavation of dental caries to ascertain the optimum depth of excavation.

Participants

Participants asymptomatic for dental caries who may have early, undetected caries at the point of recruitment. Seemingly asymptomatic patients may still have early caries which are undetected at the point of recruitment. Studies that explicitly recruited participants with caries into dentine or frank cavitation were excluded. As were those with participants referred to secondary care for restorative treatment, as there is a likelihood that advanced caries (into dentine or pulp) would be present and readily detectable without the need for the index tests investigated in this review.

Studies recruiting children, adolescents, and adults were all eligible for inclusion, this allowed for the analysis of the diagnostic test accuracy of index tests for primary, mixed, and permanent dentition.

Index tests

Transillumination and optical coherence tomography (OCT) devices that utilised white light scattering and near-infrared (NIR) to aid the detection or diagnosis of coronal caries. These index tests must be completed on intact teeth and could be used as an adjunct or replacement for aspects of those currently used in general dental examinations (e.g. transillumination or OCT as a replacement for radiography). The intention was to assess the index tests in isolation wherever possible otherwise the result of one index test may influence another. However, where multiple index tests were used as a combined index test we planned to report these separately.

Where studies used multiple examiners the most appropriate examiner to the research question were selected. For example, if the study used dental students, general dental practitioners, and restorative consultants, then the results of the general dental practitioner were extracted. In the scenario where multiple examiners were stated to have similar skills and experience, then the mean sensitivity and specificity values were extracted if available, otherwise the first set of reported results was extracted.

Target conditions

Coronal caries: initial stage decay, defined as initial or incipient caries or non-cavitated lesions. Specifically where there is a detectable change in enamel evident which is not thought to have progressed into dentine; on i) occlusal, ii) approximal surfaces, or iii) smooth surfaces.

Reference standards

A number of different reference standards have been used in primary diagnostic test accuracy (DTA) studies. The only way of achieving a true diagnosis of caries presence and depth is to extract and section the tooth and perform a histological assessment (Downer 1975; Kidd 2004). However, this method would not be ethically reasonable to undertake on a healthy population in clinical (in vivo) studies, whilst in vitro studies commonly use histology on previously extracted teeth. The only scenario where histology can be a viable scenario for studies undertaken in a primary or secondary care setting would be where a tooth has been identified as requiring extraction (ideally for a non-caries related reason, such as orthodontic extraction or third molar extraction) and the index test could be applied prior to extraction, followed by the reference standard of histology. This may bring into question

the study's broader external validity as these types of studies are most likely to occur in adolescents or younger adults who may have a lower prevalence of disease than an adult population and may not be representative of the wider population.

The optimum reference standard was histology. Alternative, acceptable reference standards for this review include operative exploration, where a clinician removes caries with a dental burr (drill) in preparation for a restoration and reports the depth of decay. This technique would be acceptable as a reference standard for patients with caries requiring restoration, but would not be ethical for caries-free patients and a different reference standard would be required, such as a radiograph, although there is concern regarding the accuracy of radiographs to detect early enamel lesions and this was accounted for in the quality assessment.

Tooth separation using orthodontic bands would be an acceptable reference standard for studies using approximal tooth surfaces. This method inserts a band between teeth which is generally left in situ for up to a week, forming a gap between the teeth and allowing a direct view of the approximal surface of the tooth which would otherwise have been obscured by the abutting tooth. Although this method of visualising the proximal surfaces is not often used in general practice this does make in vivo studies more feasible.

A period of up to three months between the index test and a reference standard was deemed acceptable.

Search methods for identification of studies

For the planned reviews on the detection and diagnosis of caries, separate search strategies were developed for MEDLINE Ovid and Embase Ovid, according to the guidance provided in Chapter 7 of the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (de Vet 2008).

Electronic searches

Cochrane Oral Health's Information Specialist conducted systematic searches in the following databases without language or publication status restrictions:

- MEDLINE Ovid (1946 to 15 February 2019) ([Appendix 1](#));
- Embase Ovid (1980 to 15 February 2019) ([Appendix 2](#)).

Searching other resources

The following trial registries were searched for ongoing studies:

- US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (clinicaltrials.gov; searched 15 February 2019) ([Appendix 3](#));
- World Health Organization International Clinical Trials Registry Platform (apps.who.int/trialsearch; searched 15 February 2019) ([Appendix 4](#)).

We searched the reference lists of included papers and previously published systematic reviews for additional publications not identified in the electronic searches.

Data collection and analysis

Selection of studies

Two review authors independently screened and assessed the results of all searches for inclusion. Any disagreements were

resolved through discussion and, where necessary, consultation with another clinical or methodological member of the author team. During the screening process, studies eligible for all reviews were identified for inclusion. Studies were excluded if they failed to present the data in the format of a 2 x 2 contingency table or failed to report prevalence at the enamel threshold to allow a computation of the 2 x 2 table. In these instances the study authors were contacted and the required data requested. An adapted PRISMA flowchart was used to report the study selection process (McInnes 2018). Once agreement for inclusion was reached, the studies were categorised according to their index test - specifically the type of illumination device, tooth surface, and age (adult or child).

Data extraction and management

Two review authors extracted data independently and in duplicate. A piloted study data extraction form based on the review inclusion criteria was developed and applied to eligible studies. Disagreements were resolved through discussion by the review team. Where data had been reported for occlusal and approximal surfaces, data were extracted separately for the different surfaces. Study authors were contacted to obtain missing data or characteristics which were not evident in the published paper.

We recorded the following data for each study:

- sample characteristics (age, sex, socioeconomic status, risk factors where stated, number of patients/carious lesions, lesion location, disease prevalence - at enamel and dentine thresholds);
- setting (country, disease prevalence, type of facility);
- the type of index test(s) used (category/scale, name, conditions (i.e. clean/dried teeth), positivity threshold);
- study information (design, reference standard, case definition, training and calibration of personnel);
- study results (true positive, true negative, false positive, false negative, any equivocal results, withdrawal).

Assessment of methodological quality

We used Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) to assess the risk of bias and applicability of the eligible primary studies over the four domains of participant selection, index test, reference standard, and flow and timing (Whiting 2011), tailored for this review. 'Review specific' descriptions of how the QUADAS-2 items were contextualised and implemented are detailed in the accompanying checklist (Additional Table 3).

A 'Risk of bias' judgement ('high', 'low', or 'unclear') was made for each domain. Generally, where the answers to all signalling questions within a domain were judged as 'yes' (indicating low risk of bias for each question) then the domain was judged to be at low risk of bias. If any signalling question was judged as 'no', indicating a high risk of bias, the domain was scored as high risk of bias. This was followed by a judgement about concerns regarding applicability for the participant selection, index test, and reference standard domains. Results of the assessment of methodological quality are presented graphically.

Participant selection domain (1)

The selection of patients has a fundamental effect on an index test's ability to detect caries. It was essential that the disease stages of

sound and enamel caries be represented in the sample and the age range of patients needed to be reported to form a complete appraisal of the test's potential to correctly classify disease in different populations.

It was acceptable for studies to focus on one particular surface (occlusal/approximal) or particular age group (children/adults). Given that the primary objective centred on early enamel lesions studies should be reporting on this stage of the disease process. It was vital that within the chosen population all participants or teeth meeting the eligibility criteria should be provided with an equal or random opportunity to be included. Inappropriate exclusion may lead to an over or under estimation of the test's ability to detect disease, thus affecting the internal validity of the study.

All studies should have clearly reported the methods used to select teeth, ideally a random or consecutive selection would be used and the procedure clearly reported. In addition it should be clearly reported which disease stages were included in the study sample and the prevalence of each stage of disease should be reported. This was used to inform the applicability of this test to a wider population.

Study results should be reported at a tooth or surface level, as opposed to patient level, which has the potential for the index test and reference standard to be reporting on different sites within the same mouth.

Index test domain (2)

The nature of the index tests and the visual presentation of the disease means that it should be feasible to ensure that the index test is conducted prior to the reference standard. The transillumination or OCT examination should be completed before the extraction of a tooth for any histological analysis, tooth separation or before in situ excavation of a tooth is undertaken. To minimise potential for bias, separate examiners should have been utilised for index test and reference standard. The threshold of disease positive and negative should have been presented prior to analysis and be reflective of the participants recruited to the study. With the studies investigating asymptomatic patients at a screening level, then early stages of disease would be of primary importance and thresholds of caries into enamel of greater relevance than caries into dentine or pulp. With the subjective nature of the index tests there may be potential for information bias unless different examiners have been applied for each of the thresholds interpreted within the studies. For example, if the decision is borderline between caries into enamel and dentine, the interpretation of the first threshold would influence the decision made on the second threshold. It is unlikely that studies will have utilised multiple index test examiners or where they have, it is probable that they each score all of the thresholds and are included for validation of the test. However, the inclusion of a question here allowed the identification of studies that have achieved this.

Reference standard domain (3)

If the reference standard was a radiograph, excavation, or tooth separation to allow a visual assessment of the approximal surface then it should have been completed by an examiner different to the index test, as the subjectivity of this type of reference standard could be compromised by knowledge of the index test results. An exception was built in for this signalling question because where the tooth had been extracted, sectioned, and prepared for

histological evaluation it is extremely unlikely that the examiner would be able to recall the specific tooth or participant and the results from the index test results. Time delays between index test and reference standard should be under three months for in vivo studies.

Ideally, each participating tooth or patient within a study should be exposed to the same reference test. This is possible in the in vitro setting as each selected tooth can have a histological assessment applied. In vivo studies may have applied the same reference standard by using enhanced visual examination or radiograph to all participants. If a study allocated participants or specific teeth to different reference tests then reasons for this allocation should have been clearly reported. All reference standards should have been completed without knowledge of the index test results.

Flow and timing domain (4)

The index test should have been conducted prior to the reference standard. If the reference standard used was tooth separation, radiograph, or excavation then there should be less than three months between index test and reference standard. Caries is a slow growing disease so minimal changes should be experienced within this time frame. All included teeth in the sample should receive both an index test and reference standard. Where studies report some teeth having an index test but not a reference standard, a reason should be clearly reported, such as teeth being broken during sectioning, this would not influence the risk of bias decision.

Comparative domain

If comparative test studies were identified and included then a comparative domain would have been added to the QUADAS-2 checklist. Selection bias needs to be considered regarding selection of teeth or participants for inclusion in between-person comparison studies (RCTs), i.e. were the same participant selection criteria used for those allocated to each test? Further considerations for studies where index tests have been compared, either direct within-person or between-person comparisons, would be the ordering of index tests and the blinding of examiners to prior or subsequent index tests. For between-person comparison studies (RCTs) there must be a maximum time delay between tests of three months, to ensure that the disease has not progressed and invalidated the comparison.

Statistical analysis and data synthesis

The threshold of interest was between sound teeth and early enamel caries. Estimates of diagnostic accuracy were expressed as sensitivity and specificity with 95% confidence intervals for each study and for each available data point if there were multiple index tests, teeth (primary/permanent) or surfaces (occlusal/approximal/smooth) reported within the same study. When there were two or more test results reported in the same study, for example the occlusal and proximal surfaces, we included them in the analysis as separate datasets. For both the overall analysis of all included datasets and the analysis of different imaging modalities we also indicated the 95% prediction regions as an indication of the region in which the sensitivity and specificity of a future study could be expected to lie given the results of the studies that have already been observed and included in the analysis. Diagnostic odds ratios (DOR), defined as the ratio of the odds of test positive in those who have the target condition relative to the odds of test positive in

those without the condition, were calculated as a global measure of accuracy.

Hierarchical models were used for data synthesis. Study estimates of sensitivity and specificity were plotted on coupled forest plots and in receiver operating characteristic (ROC) space. Meta-analysis was conducted which combined the results of studies for each index test using the bivariate approach to estimate the summary values of sensitivity and specificity at a common threshold (Macaskill 2010; Reitsma 2005). Data were input into Review Manager 5 (Review Manager 2020) and displayed as coupled forest plots. The bivariate binomial method was applied by fitting a generalized linear mixed model (GLMM) using the glmer function in the R package lme4 (Bates 2007; Partlett 2016), along with the MetaDTA interactive web-based tool (Freeman 2019). We used a meta-regression approach to compare the accuracy of the different device types in this review. We added device type as covariates to the bivariate model, assuming equal variances, and used a likelihood ratio test to formally assess the significance of any model comparisons. Initially we allowed test type to be assessed on both sensitivity and specificity. If a difference in sensitivity or specificity or both was observed then further investigations were undertaken to determine whether the differences could be attributed to sensitivity or specificity.

Investigations of heterogeneity

Sources of heterogeneity were considered individually. Initially, an inspection of the clinical and methodological characteristics of the included studies, coupled forest plots and summary ROC plots were used to form the basis of the assessment of heterogeneity. More formally, meta-regression analyses were carried out to explore possible sources of heterogeneity, where sufficient numbers of studies allowed (Rutter 2001). Formal model comparisons were undertaken using likelihood ratio χ^2 tests to determine the statistical significance of adding one or more potential sources of heterogeneity (covariates) to the bivariate model. Where substantial heterogeneity was observed then this was clearly articulated in the 'Summary of findings' tables to aid interpretation of the results.

The sources of heterogeneity included (specified a priori).

- Population characteristics.
 - Children or adults; the detection of disease in the primary or permanent dentition.
 - In vitro or in vivo studies; the accuracy estimates from laboratory-based studies conducted on extracted teeth could conceivably be higher than those of clinical-based studies that are subject to the additional complexities of detection and diagnosis within the oral cavity in the presence of plaque and staining of a tooth's surface. Further, in vitro studies typically employ a reference standard of histology whereas in vivo studies may employ excavation or an enhanced visual examination with radiographs for this purpose.
 - Selection of tooth surface under investigation.
 - Participants or teeth with previously applied pit and fissure sealants.
 - Prevalence of caries into dentine. We created three categories for the prevalence of dentine caries: low $\leq 14\%$, medium 15% to 34% , and high $\geq 35\%$. As studies that explicitly included

observations with more advanced lesions into dentine or frankly cavitated or lesions that were frankly cavitated were excluded, we anticipated that the prevalence of caries into dentine of the included studies would typically be low.

- Index test characteristics.
- Different illumination devices.

Sensitivity analyses

Where a sufficient number of studies investigated the same index test, the following sensitivity analyses were performed to assess the impact on summary estimates of restricting the analyses according to the following criteria:

- low prevalence of dentine caries (i.e. less than 15%);
- low risk of bias for an index test;
- low risk of bias exists for a reference standard.

Assessment of reporting bias

Methods currently available to assess reporting or publication bias for diagnostic studies may lead to uncertainty and misleading results from funnel plots (Deeks 2005; Leeflang 2008), therefore we did not perform reporting bias tests in the reviews.

Summary of findings and assessment of the certainty of the evidence

We reported our results for illumination index tests and for the main target conditions following GRADE methods (Schünemann

2020; Schünemann 2020a), and using the GRADEPro online tool (www.guidelinedevelopment.org). To enhance readability and understanding, we presented test accuracy results in natural frequencies to indicate numbers of false positives and false negatives. The certainty of the body of evidence was assessed with reference to the overall risk of bias of the included studies, the directness of the evidence, the inconsistency of the results, the precision of the estimates, and the risk of publication bias; these have been considered narratively where statistical methods were not available. We categorised the certainty of the body of evidence as high, moderate, low, or very low.

RESULTS

Results of the search

The search retrieved 1091 results, 718 remained after the removal of duplicates. The initial screening of titles and abstracts resulted in 70 studies being considered for inclusion. We obtained full papers for these studies and after further review of the full text we included 23 and excluded 47 studies. The most common reason for exclusion was that the studies failed to present the data in a format that allowed us to create a 2 x 2 table of the results at the enamel threshold (28 studies) or studies intentionally included frank cavitation or dentinal lesions (14 studies). Other reasons for exclusion were a lack of reference standard (one study), an inappropriately applied reference standard such as visual examination being applied to non-separated approximal surfaces (one study), or the exclusion of sound teeth (three studies) (Figure 2).

Figure 2. Study flow diagram.

We included 23 studies in this review. 16 studies investigated a second index test which is detailed in another review of this series. Of these, six also investigated radiographs (Astvaldsdottir 2012; Jablonski-Momeni 2017; Mialhe 2003; NCT02657538; Shimada 2014; Simon 2016), another six reported the results of a visual examination (Laitala 2017; Nakagawa 2013; Nakajima 2014; Shimada 2010; Sidi 1988; Xiao-Hua 2016), and two reported a fluorescence device (Mansour 2016; Van Hilsen 2013). Two studies reported multiple index tests (Ashley 1998; Bussanelli 2015). No study reported two different types of transillumination or optical coherence tomography (OCT) device within the same study.

Four studies were completed in the United States (Holtzman 2010; Mansour 2016; Simon 2016; Van Hilsen 2013) and Japan (Nakagawa 2013; Nakajima 2014; Shimada 2010; Shimada 2014), three studies in Germany (Jablonski-Momeni 2017; Lederer 2018; Park 2018), and two in the UK (Ashley 1998; Holtzman 2010) and Brazil (Bussanelli 2015; Mialhe 2003). The remainder were from other European countries: Sweden (Astvaldsdottir 2012), Switzerland (NCT02657538), Finland (Laitala 2017), and France (Obry-Musset 1988). Also there was one study from each of Australia (Chawla 2012), China (Xiao-Hua 2016), and Malaysia (Zain 2018); one study did not report its country of origin (Sidi 1988).

The included studies date from 1988 to 2019. Five of the fibre-optic transillumination (FOTI) based studies were completed prior to 2005 (Ashley 1998; Holt 1989; Mialhe 2003; Obry-Musset 1988; Sidi 1988), with digital fibre-optic transillumination (DIFOTI) being investigated more recently (Astvaldsdottir 2012; Laitala 2017). All OCT studies have been completed since 2010 (Holtzman 2010; Mansour 2016; Nakagawa 2013; Nakajima 2014; Park 2018; Shimada 2010; Shimada 2014; Van Hilsen 2013; Xiao-Hua 2016; Zain 2018), whereas the near-infrared (NIR) devices have all been studied since 2015 (Bussanelli 2015; Jablonski-Momeni 2017; Lederer 2018; NCT02657538; Simon 2016).

In total there were six FOTI studies (Ashley 1998; Chawla 2012; Holt 1989; Mialhe 2003; Obry-Musset 1988; Sidi 1988) and two DIFOTI (Astvaldsdottir 2012; Laitala 2017). 10 studies investigated an OCT device (Holtzman 2010; Mansour 2016; Nakagawa 2013; Nakajima 2014; Park 2018; Shimada 2010; Shimada 2014; Van Hilsen 2013; Xiao-Hua 2016; Zain 2018). Of these studies seven were described as swept-source OCT (Mansour 2016; Nakagawa 2013; Nakajima 2014; Shimada 2010; Shimada 2014; Xiao-Hua 2016; Zain 2018), two spectral-domain (Holtzman 2010; Park 2018), and one as cross-polarisation (Van Hilsen 2013). There were five NIR transillumination studies (Bussanelli 2015; Jablonski-Momeni 2017; Lederer 2018; NCT02657538; Simon 2016), with Simon 2016a being included as a duplicate study since it investigated occlusal and approximal surfaces. Of the NIR device studies, two investigated DIAGNOcam (Lederer 2018; NCT02657538) and one VistaCam (Jablonski-Momeni 2017), the other two studies investigated prototype devices (Bussanelli 2015; Simon 2016).

There were a total of 24 datasets available for analysis from the 23 included studies. From this point onwards we will refer to the datasets rather than individual studies. 12 datasets were completed on occlusal surfaces (Astvaldsdottir 2012; Chawla 2012; Holt 1989; Jablonski-Momeni 2017; Laitala 2017; Lederer 2018; Mialhe 2003; NCT02657538; Obry-Musset 1988; Shimada 2014; Sidi 1988; Simon 2016a) and nine used devices to investigate the approximal surface (Ashley 1998; Bussanelli 2015; Holtzman 2010; Nakajima 2014; Shimada 2010; Simon 2016; Van Hilsen 2013; Xiao-Hua 2016; Zain 2018). Of the remaining studies Mansour 2016 investigated all available surfaces of the tooth and did not separate the results according to surface and two studies reported on smooth surfaces (Nakagawa 2013; Park 2018). Three studies focused on primary dentition (Chawla 2012; Holt 1989; Nakajima 2014), 18 reported on permanent teeth (Ashley 1998; Astvaldsdottir 2012; Bussanelli 2015; Jablonski-Momeni 2017; Laitala 2017; Lederer 2018; Mansour 2016; Mialhe 2003; NCT02657538; Park 2018; Shimada 2010; Shimada 2014; Sidi 1988; Simon 2016; Simon 2016a; Van Hilsen 2013; Xiao-Hua 2016; Zain 2018), one reported on a mixture of primary and permanent (Obry-Musset 1988) although only permanent results are included in our analysis, and two did not report which type of tooth was under investigation (Holtzman 2010; Nakagawa 2013).

Thirteen datasets were completed in vitro on previously extracted teeth (Ashley 1998; Astvaldsdottir 2012; Bussanelli 2015; Chawla 2012; Holtzman 2010; Lederer 2018; Nakagawa 2013; Nakajima 2014; Park 2018; Shimada 2010; Van Hilsen 2013; Xiao-Hua 2016; Zain 2018) and all of these studies used histology as the reference standard, apart from two (Lederer 2018; Park 2018) which used microcomputed tomography (microCT) and a visual examination, respectively. The in vivo studies used various reference standards including visual with separation (Jablonski-Momeni 2017; Mialhe 2003; NCT02657538), a combination of visual and radiograph (Mansour 2016), and visual examination combined with excavation of more severely affected surfaces (Shimada 2014). Four studies used radiographs (Holt 1989; Laitala 2017; Obry-Musset 1988; Sidi 1988) and one used histology after the index test had been completed in a clinical setting and the teeth were subsequently extracted (Simon 2016; Simon 2016a).

Methodological quality of included studies

None of the included studies can be considered as being at low risk of bias across all domains of the QUADAS-2 criteria (Figure 3; Figure 4). Patient selection had no studies that were at low risk of bias, with 70% being at high risk. The index tests were generally well conducted with only 5% being at high risk of bias but the reference standard domain saw this rise to 43%. Applicability was a concern in the patient selection domain (61% of studies), but only 17% for the index test and reference standard domains.

Figure 3. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.

Figure 4. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Ashley 1998	+	+	+	+	+	+	+
Astvaldsdottir 2012	+	?	+	+	+	+	+
Bussanelli 2015	+	?	+	+	+	+	+
Chawla 2012	+	+	+	+	+	+	+
Holt 1989	?	+	+	+	+	+	+
Holtzman 2010	+	+	+	+	+	+	+
Jablonski-Momeni 2017	?	+	+	+	+	+	+
Laitala 2017	?	?	+	+	+	+	+
Lederer 2018	+	?	?	+	+	+	+
Mansour 2016	?	+	+	+	+	+	+
Mialhe 2003	?	+	+	+	+	+	+
Nakagawa 2013	+	+	+	+	+	+	+
Nakajima 2014	+	+	+	+	+	+	+
NCT02657538	+	+	+	+	+	+	+
Obry-Musset 1988	?	+	+	+	+	+	+
Park 2018	+	?	+	+	+	+	+
Shimada 2010	+	+	+	+	+	+	+
Shimada 2014	+	+	+	+	+	+	+
Sidi 1988	+	?	+	?	+	+	+
Simon 2016	?	+	+	+	+	+	+
Simon 2016a							
Van Hilsen 2013	+	+	+	+	+	+	+
Xiao-Hua 2016	+	+	+	+	+	+	+
Zain 2018	+	+	+	+	+	+	+

+ High
 ? Unclear
 + Low

Figure 4. (Continued)

Patient selection was at high risk of bias in 16 studies which clearly reported that participants or teeth were selected rather than being recruited consecutively or randomly (Ashley 1998; Astvaldsdottir 2012; Bussanelli 2015; Chawla 2012; Holtzman 2010; Lederer 2018; Nakagawa 2013; Nakajima 2014; NCT02657538; Park 2018; Shimada 2010; Shimada 2014; Sidi 1988; Van Hilsen 2013; Xiao-Hua 2016; Zain 2018). The concern here is that teeth or participants may have been selected with caries that were more difficult or more straightforward to detect which would introduce bias and influence sensitivity and specificity. Seven studies did not clearly report how participants were selected, these were recorded as being at unclear risk of bias (Holt 1989; Jablonski-Momeni 2017; Laitala 2017; Mansour 2016; Mialhe 2003; Obry-Musset 1988; Simon 2016).

Nine studies were observed as being at low concern regarding the applicability of the participant selection, this is due to them conducting the examination in a clinical setting (Holt 1989; Jablonski-Momeni 2017; Laitala 2017; Mansour 2016; Mialhe 2003; NCT02657538; Obry-Musset 1988; Shimada 2010; Simon 2016). Of the remaining studies 13 were at high concern as they investigated previously extracted teeth and Sidi 1988 was at high concern because the included participants had an extremely low prevalence of both enamel and dentinal caries which is not generalisable to the general population.

Sixteen studies were at low risk of bias for the index test domain. Unclear decisions were reached because: we could not be certain that the index test device thresholds were predetermined in one study (Bussanelli 2015), the examiners may have been influenced by other index tests (Astvaldsdottir 2012; Lederer 2018; Sidi 1988), or there may have been a lack of independence during examination of the index test and the reference standard (Laitala 2017; Park 2018). One study was at high risk as the same examiner completed two index tests and the results from a fluorescence test may have influenced the results (Mansour 2016). Applicability was seen to be of high concern in four studies as they used a device (Bussanelli 2015; Mansour 2016; Simon 2016) or technique (Lederer 2018) which is not currently available to the general dental practitioner.

The reference standard domain showed 12 studies to be at low risk of bias, with one being unclear (Lederer 2018) as we were uncertain whether the same examiner was used for index test and reference standard. 10 studies were at high risk as we could not be certain that the reference standard would provide the correct results since they used a visual examination (Jablonski-Momeni 2017;

Mansour 2016; Mialhe 2003; NCT02657538; Park 2018; Shimada 2014) or radiographs (Holt 1989; Laitala 2017; Obry-Musset 1988; Sidi 1988). These four radiograph studies were also of high concern for applicability as the results of the radiograph review in this series confirm that radiographs are not accurate at identifying early enamel caries so are not an applicable reference standard.

Six studies were found to be at high risk of bias for the flow and timing domain, this was because they failed to include all participants in the final results (Bussanelli 2015; Holt 1989; Mialhe 2003; Obry-Musset 1988; Shimada 2014; Zain 2018), one study was unclear regarding the number of missing participants in the results (Sidi 1988).

Findings

We evaluated the accuracy of transillumination and OCT devices in 23 studies with 24 datasets being included in the meta-analysis, one study used an NIR device to assess occlusal and approximal tooth surfaces so a second dataset was generated (Simon 2016; Simon 2016a). Transillumination and OCT devices were used to detect early/enamel caries in 16,702 tooth sites or surfaces, with an overall prevalence of enamel caries of 15%. Nine datasets reported at the tooth level (Ashley 1998; Bussanelli 2015; Chawla 2012; Holtzman 2010; Mansour 2016; Mialhe 2003; Simon 2016; Van Hilsen 2013; Xiao-Hua 2016), 10 datasets reported multiple surfaces of a tooth - these were all investigating approximal surfaces (Astvaldsdottir 2012; Holt 1989; Jablonski-Momeni 2017; Laitala 2017; Lederer 2018; NCT02657538; Obry-Musset 1988; Shimada 2014; Sidi 1988; Simon 2016a), and the remaining five datasets reported multiple sites per surface - on either occlusal (Nakajima 2014; Park 2018; Shimada 2010; Zain 2018) or smooth surfaces (Nakagawa 2013).

Figure 5 and Figure 6 present the main study results which are also reported in the Summary of findings 1. The primary findings are reported for all available datasets with no restrictions on the type of device used, the tooth surface, dentition, reference standard, or prevalence of disease. The sensitivity results for individual studies ranged from 0.21 to 1.00 and the specificities from 0.50 to 0.98. The analysis was undertaken using a bivariate model, the estimate of summary sensitivity and specificity points were 0.75 (95% confidence interval (CI) 0.62 to 0.85) and 0.87 (95% CI 0.82 to 0.92). The diagnostic odds ratio (DOR) was 21.52 (95% CI 10.89 to 42.48).

Figure 5. Forest plot of all included datasets.

Figure 6. Summary receiver operating characteristic (SROC) plot of all included datasets.

There were three main groups of devices that used differing technologies so it was important to investigate the results of these as subgroups. The studies are subgrouped in [Figure 7](#) and the

resulting summary receiver operating characteristic (SROC) plot can be seen in [Figure 8](#).

Figure 7. Forest plot of tests subgroups: near-infrared (NIR), optical coherence tomography (OCT), and fibre-optic transillumination (FOTI)/digital fibre-optic transillumination (DIFOTI).

NIR

OCT

FOTI/DIFOTI

Figure 8. Summary receiver operating characteristic (SROC) plot of tests highlighting the 3 different groups (near-infrared (NIR), optical coherence tomography (OCT), and fibre-optic transillumination (FOTI)/digital fibre-optic transillumination (DIFOTI)).

- FOTI and DIFOTI: eight included datasets had an estimated summary point for sensitivity of 0.47 (95% CI 0.35 to 0.59) and specificity 0.92 (95% CI 0.86 to 0.96).
- OCT: 10 datasets reported an estimated summary point for sensitivity of 0.94 (95% CI 0.88 to 0.97) and specificity 0.83 (95% CI 0.68 to 0.91). [Figure 9](#) is included to highlight the high levels of uncertainty between studies included in the OCT group.
- NIR: six datasets reported an estimated summary point for sensitivity of 0.58 (95% CI 0.46 to 0.68) and specificity 0.86 (95% CI 0.80 to 0.91).

Figure 9. Illustration of the uncertainty in the optical coherence tomography (OCT) devices.

A formal comparison of the accuracy according to device type indicated a difference in sensitivity or specificity or both ($\text{Chi}^2(4) = 34.17$, $P < 0.01$). Further analysis indicated a difference in the sensitivity of the different devices ($\text{Chi}^2(2) = 31.24$, $P < 0.01$) with a higher sensitivity of 0.94 (95% CI 0.88 to 0.97) for OCT compared to NIR 0.58 (95% CI 0.46 to 0.68) and FOTI/ DIFOTI 0.47 (95% CI 0.35 to 0.59), but no meaningful difference in specificity ($\text{Chi}^2(2) = 3.47$, $P = 0.18$).

Investigations of heterogeneity

Investigations of heterogeneity could not be conducted using a meta-regression for the combined group of all studies due to the variation observed between device groups. The potential areas of heterogeneity could be investigated across the three groups (NIR,

OCT, and FOTI/DIFOTI), as a narrative and also by plotting variations in SROC space and making observations.

Dentition

Eighteen of the 24 datasets (75%) investigated the transillumination and OCT devices on permanent teeth. Due to this high proportion of one group being reported the heterogeneity could not be investigated. The subgroups investigated dentition as follows:

- NIR: all six datasets reported on permanent teeth;
- OCT: seven of the 10 datasets reported on permanent teeth;
- FOTI/DIFOTI: five of the eight datasets investigated permanent teeth, two primary, and one was mixed.

Prevalence of dentine lesions

Two of the device subgroups did not have sufficient studies across all levels of prevalence to allow for an investigation of this covariate: NIR - three low and three medium; FOTI/DIFOTI - one high, four medium, and three low.

The OCT group had two high, three medium, and five low prevalence. Due to this range of included prevalences at the dentine

level these were investigated in the forest plots (Figure 10) and plotted in SROC space (Figure 11). It can be seen from the SROC curve that the datasets with higher prevalence of caries into dentine have higher sensitivity and specificity: the two studies with a prevalence > 35% both have a sensitivity and specificity greater than 0.9 (Nakagawa 2013; Nakajima 2014). This should be interpreted with caution as one other study reported similar levels of sensitivity and specificity but had a prevalence of dentine caries of zero (Zain 2018).

Figure 10. Forest plot of optical coherence tomography (OCT) datasets investigating the effect of prevalence at the dentine level.

Figure 11. Summary receiver operating characteristic (SROC) plot of optical coherence tomography (OCT) datasets, highlighted according to the prevalence of dentine level caries.

Surface

It was not possible to investigate whether the tooth surface assessed (occlusal, approximal, or smooth) had an effect on the sensitivity and specificity results. Each subgroup was heavily weighted to one type of surface which prevented the analysis of this covariate. The surfaces investigated within each subgroup were:

- NIR: four studies approximal and two occlusal;
- OCT group: six occlusal, one approximal, one mixed, and two smooth;
- FOTI/DIFOTI: seven of the eight studies investigated approximal surfaces.

In vivo or in vitro studies with different reference standards

The reference standard applied was able to be investigated for each of the subgroups, through the observation of forest plots and SROC plots.

- NIR had four studies which used either histology or microcomputed tomography (microCT) which resulted in a low risk of bias for the reference standard domain of QUADAS-2 but also relied on the tooth being extracted. Two studies used a visual reference standard. When viewed on forest plots and

SROC plots there was no observable effect on the sensitivity and specificity results for this subgroup (Figure 12; Figure 13).

- OCT had seven of the 10 datasets using histology and the other three using a visual examination. It appears from the forest plots (Figure 14) and SROC plot (Figure 15) that those with histological reference standards produced results with higher sensitivity and specificity.
- FOTI/DIFOTI had three studies that used histology, four radiographs, and one visual examination as a reference standard. There was no observable effect on the sensitivity and specificity results (Figure 16; Figure 17).

Figure 12. Forest plot of 2 near-infrared (NIR).

Figure 13. Summary receiver operating characteristic (SROC) plot of the near-infrared (NIR) group, individual datasets being highlighted according to the reference standard applied.

Figure 14. Forest plot of optical coherence tomography (OCT) sorted according to the reference standard applied.

Figure 15. Summary receiver operating characteristic (SROC) plot of the optical coherence tomography (OCT) group, individual datasets being highlighted according to the reference standard applied.

Figure 16. Forest plot of the fibre-optic transillumination (FOTI)/digital fibre-optic transillumination (DIFOTI) group, sorted according to reference standard applied.

Figure 17. Summary receiver operating characteristic (SROC) plot of the fibre-optic transillumination (FOTI)/digital fibre-optic transillumination (DIFOTI) group, individual datasets being highlighted according to the reference standard applied.

Multiple or single sites of assessment

We were interested to investigate whether any of the subgroups was affected by the use of multiple sites per tooth.

- NIR: used a tooth level investigation in two studies and multiple surfaces on the approximal surfaces.
- FOTI/DIFOTI: investigated either the whole tooth (three datasets) or multiple surfaces (five datasets), however these multiple surfaces were approximal surfaces which naturally provide two surface per tooth which are likely to be independent of each other.
- OCT: there were four datasets that reported at the tooth level and one at the surface level for approximal studies, however

the remaining five studies all investigated multiple sites per tooth for occlusal or smooth surfaces, with a high potential for contamination between these sites. These were investigated as a covariate across two groups, those with multiple sites on a single surface and the tooth level and approximal surface

datasets (Figure 18). The estimate of summary sensitivity was higher for the multiple sites group (0.97 (95% CI 0.94 to 0.98)) than the single site group (0.86 (95% CI 0.78 to 0.92)) and the specificity for the multiple group (0.87 (95% CI 0.68 to 0.96)) was higher than the single group (0.77 (95% CI 0.54 to 0.90)).

Figure 18. Summary receiver operating characteristic (SROC) plot of optical coherence tomography (OCT) studies illustrating those that assessed a single site on the tooth's surface and those that assessed multiple sites on a tooth's surface.

Sensitivity analyses were planned to assess levels of bias across the QUADAS-2 domains. The first sensitivity analysis suggested

we explore the effect of unclear reporting of participant selection, this occurred in six studies (Holtzman 2010; Laitala 2017; Mialhe

2003; Obry-Musset 1988; Simon 2016) these were spread across the different subgroups of OCT (two studies), NIR (one study), and FOTI (three studies). We excluded the studies at unclear risk of selection bias from the primary analysis and this did not alter our estimates of sensitivity and specificity. The second sensitivity analysis planned to investigate high prevalence which has been investigated above. The third sensitivity analysis was to assess the effect of a low risk of bias in the index test domain, this occurred in all studies except one, so was not necessary. The final sensitivity analysis assessed the effect of bias on the reference standard, we looked at this for studies with high risk of bias, which occurred in 10 studies (four OCT, four FOTI, and two NIR) (Holt 1989; Jablonski-Momeni 2017; Laitala 2017; Mansour 2016; Mialhe 2003; NCT02657538; Obry-Musset 1988; Park 2018; Shimada 2010; Sidi 1988). The results of the sensitivity analysis for OCT were a sensitivity of 0.94 (95% CI 0.88 to 0.97) and specificity 0.84 (95% CI 0.67 to 0.93); FOTI sensitivity 0.47 (95% CI 0.27 to 0.67) and specificity 0.94 (95% CI 0.90 to 0.96); none of which were of substantial difference to the summary measures reported in the main analyses.

DISCUSSION

Summary of main results

The included studies allowed us to evaluate the diagnostic test accuracy of transillumination and optical coherence tomography (OCT) devices for the detection of caries, with a particular focus on early-stage caries in the enamel of the tooth. Specifically, we estimated the accuracy of fibre-optic transillumination using fibre-optic transillumination (FOTI) or digital fibre-optic transillumination (DIFOTI), near-infrared (NIR) transillumination, and OCT. Sufficient studies enabled the construction of 2 x 2 tables, and we were able to conduct a meta-analysis. The included studies covered three categories of illumination devices that exploit various methods of application and interpretation, each primarily defined by different wavelengths. When we explored the diagnostic test accuracy within these subgroups, our conclusions became more limited.

We observed variation in the estimated summary sensitivity and specificity for each device category, and the included studies were diverse in their study design, sample population, index test, and reference standard. This clinical and methodological heterogeneity should be considered in the interpretation of the results of this review. The low methodological quality of the available studies was due in part to unavoidable difficulties in study design, and we were unable to categorise any study as low risk of bias across all domains and as low concern for applicability.

The participant selection domain had the largest number of studies judged to be at high risk of bias (16 studies). The study sample should be recruited either consecutively or randomly to avoid the suggestion that teeth included in the study sample are more complex or straightforward to diagnose. Concerns regarding the applicability of the evidence were judged as high or unclear for all studies for the patient selection domain, principally due to the large number of studies that were carried out in a laboratory setting on extracted teeth. Whilst we acknowledge this is an integral part of the research pathway, such studies inevitably elicited a judgement of high concern for applicability to our research question. The generalisability of our overall results to a clinical setting, with the inherent challenges of access to the oral cavity, patient acceptability, and time constraints is a concern. We judged

the majority of studies to be at low risk of bias and applicability concern for the index test, reference standard, and flow and timing domains. Reasons for the high risk of bias judgements for the index test included a lack of independent examiners for the index test and reference standard. We judged studies to be at a high risk of bias for the reference standard domain where an imperfect reference standard, such as a visual examination or radiograph, was used. The reference standard is inextricably linked to the study design (in vitro or in vivo), and this is a particular challenge for caries diagnostic test accuracy studies. Histology is currently the most reliable reference standard to correctly classify the target condition, and is typically carried out on extracted teeth. The use of extracted teeth, however, will automatically elicit a judgement of high concern for applicability for the participant selection domain. This challenge may be circumvented by targeting patients who required a tooth extraction for orthodontic or periodontic purposes, applying the index test to tooth surfaces within the oral cavity, and applying a histological reference standard once the teeth have been extracted. Such studies are, however, difficult to organise and administer, and could still be considered to lack broader applicability since they often focus exclusively on a particular subgroup of individuals such as adolescents who require extractions for orthodontic purposes.

We rated the certainty of the evidence as low and downgraded two levels in total due to avoidable and unavoidable study limitations in the design and conduct of studies, indirectness arising from the in vitro studies, and imprecision of the estimates.

We were particularly interested in the effects of the different illumination devices as an adjunct to a conventional visual examination. No studies formally reported the use of the devices in this way, and therefore it has not been possible to make an assessment. The search did not yield any eligible case-control or randomised controlled trials of diagnostic test accuracy. There was a limited number of studies that investigated secondary caries (around sealants or restorations), and although these study designs were defined as a potential source of heterogeneity a priori, we were unable to fully investigate in this review. One study evaluated the test accuracy of OCT on surfaces with a sealant and concluded that OCT could detect disease below a sealant (Holtzman 2010).

The main findings of this review are.

- **The estimated summary point for all included studies is sensitivity 0.75 (95% confidence interval (CI) 0.62 to 0.85) and specificity 0.87 (95% CI 0.82 to 0.92).** In a cohort of 1000 tooth surfaces with an enamel caries prevalence 57%, 142 tooth surfaces would be classified as disease-free when enamel caries was truly present (false negatives), and 56 tooth surfaces would be classified as diseased in the absence of enamel caries (false positives). The consequences of these misclassifications are concerning, and all interventions have a cost at a patient or system level. A false-positive classification for enamel caries would typically result in the application of topical fluoride or other minimally invasive treatments. A false-negative classification implies that patients who require treatment would not receive it, at that point in time. The clinician may be reassured that given the recall period for routine dental examinations and the slow-growing nature of the disease, the lesion could be identified at the patient's next appointment. The prevalence of enamel caries applied to this scenario is representative of that observed for the OCT and NIR

devices and is directly comparable to other reviews in this series. Clinicians may be concerned that this prevalence is higher than that observed in general practice, and so a lower prevalence of 28% enamel caries has also been applied. With this lower prevalence, the scenario of 1000 tooth surfaces resulted in 70 false negatives and 94 false positives ([Summary of findings 1](#)) ([Figure 19](#)).

- **There is a statistically significant difference in the accuracy of OCT, FOTI/DIFOTI, and NIR.** 42% of the available datasets allowed us to assess OCT at the level of enamel caries (10 out of 24), with 33% and 25% for FOTI/DIFOTI and NIR respectively. There was considerable heterogeneity in sensitivity and specificity within each of these groups, this resulted from different reference standards, prevalence of severe caries, tooth surfaces, and type of dentition. The full results for each group can be viewed in the [Summary of findings 1](#) and [Figure 8](#) detail the results according to the device type. Sensitivity was highest for OCT at 0.94 (95% CI 0.88 to 0.97) followed by NIR 0.58 (95% CI 0.46 to 0.68) and FOTI/DIFOTI 0.47 (95% CI 0.35 to 0.59). Specificity was similar across the devices at 0.83 (95% CI 0.68 to 0.91), 0.86 (95% CI 0.80 to 0.91) and 0.92 (95% CI 0.86 to 0.96), respectively. Interestingly, specificity was highest for FOTI/DIFOTI. This finding could be due to FOTI being easily confounded by factors such as staining or structural issues which could lower the sensitivity and increase specificity because the examiner does not want to classify the observation as caries when it truly is. The difference in accuracy across the device types meant that we were unable to explore sources of heterogeneity for all studies combined and that further analyses would have to be undertaken separately according to the device type.
- **In the OCT group, accuracy estimates were higher for studies with a high prevalence of dentinal caries ($\geq 35\%$) compared to those with a medium (15% to 34%) or low ($\leq 14\%$) prevalence, and may account for the apparent greater accuracy of OCT over FOTI/DIFOTI and NIR.** Caries into dentine is considerably more straightforward to detect than enamel caries, and it is reasonable to assume that accuracy measures could be inflated when a sample contains a high proportion of more advanced lesions. We were unable to investigate the effects of the prevalence of enamel caries for FOTI/DIFOTI and NIR devices due to insufficient studies within each of the low, medium, and high categories. This is perhaps to be expected as we restricted the inclusion criteria to studies comprising sound tooth surfaces or surfaces with initial caries. When viewed in summary receiver operating characteristic (SROC) space ([Figure 11](#)), the two studies with a high prevalence of enamel caries reported high sensitivity and specificity, whereas the results for the medium and low categories were more varied. These findings should be interpreted with caution as there were insufficient data to investigate this potential source of heterogeneity more formally.
- **There is insufficient evidence to determine whether there is a difference in accuracy between in vitro and in vivo studies.** The composition of the study sample according to extracted teeth (in vitro studies) or teeth in situ (in vivo studies) was a key consideration that informed the risk of bias and applicability assessments. Results from in vitro studies are integral to the development of devices but rarely replicate the conditions of a clinical setting. The examination of caries in situ involves difficulties in viewing all tooth surfaces within

the oral cavity, perhaps with other teeth or soft tissues obstructing the clinician's view. Moisture within a lesion is also an issue, as the refractive index of porous enamel is affected by water content. This can be an issue for in vitro studies if teeth are not stored correctly and dry out. However, the in vitro setting does facilitate the use of histology as a reliable and accurate reference standard; microcomputed tomography (microCT) has more recently been used with some confidence and was used as the reference standard in one included study ([Mansour 2016](#)). Since it is not possible to extract, section, and perform histology on healthy teeth, clinically-based studies have often use enhanced visual examination or radiographs as effectively imperfect reference standards. Differences in accuracy according to reference standard/in vitro and in vivo study design could not be formally assessed in the OCT group as only two studies applied the index test in a clinical (in vivo) setting ([Mansour 2016](#); [Shimada 2014](#)). Of the NIR studies, only two were carried out in a clinical setting ([Jablonski-Momeni 2017](#); [NCT02657538](#)). In contrast, three of the eight studies that evaluated the FOTI/DIFOTI devices used extracted teeth ([Ashley 1998](#); [Astvaldsdottir 2012](#); [Chawla 2012](#)). These findings should be interpreted with caution as there were insufficient data to investigate this potential source of heterogeneity more formally.

- **It was not possible to investigate whether the use of occlusal or approximal surfaces resulted in higher levels of accuracy.** A further confounding issue is the tooth surface that was investigated. The majority of studies were on either occlusal or approximal surfaces. The devices may have the ability to investigate approximal caries which may reduce the need for radiographs. The studies included in the meta-analysis were limited in numbers within their subgroups and it was not possible to investigate the effect of surfaces. FOTI/DIFOTI was applied to approximal surfaces in eight of the nine studies, in four of the six NIR studies but only in one of the nine OCT studies - although three more studies investigated smooth surfaces of extracted teeth which may equate to the approximal surface if the teeth were still in situ but they cannot be classed as approximal sites for the purposes of this review as the teeth were not set up in a model to replicate the abutment of teeth.
- **We were unable to investigate the effect of transillumination and OCT devices on different tooth types due to the limited number of studies of primary teeth.** Only five of the included studies evaluated the accuracy of illumination devices on the primary dentition and one study evaluated a mixture of primary and permanent teeth.
- **Studies that assessed multiple sites on a tooth's surface reported lower accuracy than those that assessed a single site.** There was concern that there may be some dependency of results from an underlying or hidden lesion when multiple sites of a tooth's surface were assessed. We were able to formally assess this hypothesis for the OCT devices with five studies detailing multiple sites, and five studies detailing single sites. The results suggested that studies that evaluated multiple sites per tooth reported higher sensitivity, although this was not statistically significant. There is some suggestion within the literature that caries detection on occlusal surfaces with OCT may be problematic and could justify multiple sites of investigation. The complex geometry of the occlusal surface may result in reflections which may result in higher sensitivity results ([Gomez 2013](#); [Zain 2018](#)).

Figure 19.

Strengths and weaknesses of the review

The strengths of this review are the completion of a comprehensive electronic literature search and rigorous application of evidence synthesis methodology that ensured that all screening, inclusion, data extraction, and quality assessment activities were carried out in duplicate and with clinical input. We employed transparent and reproducible methods for the application of methodological quality decisions. Meta-analyses provided summary estimates of sensitivity and specificity for the 23 studies overall and by device type, however, the relatively small number of included studies for each subgroup limited formal investigations of potential sources of heterogeneity according to the device type. We restricted our inclusion criteria to studies that provided sufficient data to construct a 2 x 2 table of true positives, true negatives, false positives, and false negatives. We contacted study authors where appropriate to ensure that data could be gathered for the most studies. Our review of fluorescence-based devices in this series of systematic reviews did not restrict eligibility according to this criteria and reported that of a total of 133 diagnostic test accuracy studies meeting the inclusion criteria only 79 (59%) provided useable data (Macey 2020). This finding highlights the real issue of incomplete data reporting.

Small numbers of included studies investigating FOTI, low quality of evidence, and an absence of meta-analyses have affected the utility of other systematic reviews of illumination devices (Bader 2002; Gomez 2013). The systematic review by Gomez et al also focused on early enamel lesions but, did not exclude studies that explicitly included observations with caries into dentine or frank cavitation, which may have inflated the reported accuracy of the devices. Such studies were ineligible in this review.

We initially elected to combine all studies to permit an overall meta-analysis, followed by a formal investigation of the accuracy of the device types, to allow the results of these analyses to guide the subsequent explorations of potential sources of heterogeneity. It should be noted, however, that the test accuracy comparisons were indirect comparisons, rather than direct comparisons, using for example, a paired diagnostic or a randomised design. Research has shown that evidence from indirect comparisons may differ from that of comparative studies (Takwoingi 2013), and therefore some caution is advised when interpreting the observed superior accuracy of OCT compared with NIR and fibre-optic devices.

As we observed significant differences in the accuracy of the different test types, we were unable to proceed to formal investigations of heterogeneity for all 24 datasets combined. Due to the relatively small number of studies per device, we were unable to formally investigate sources of heterogeneity according to test type, however, we have provided a narrative according to information presented in the SROC and coupled forest plots.

Applicability of findings to the review question

We have some concerns regarding the clinical applicability of the findings of this review based on the fact that 11 of the datasets in this review are the product of in vitro studies that have been conducted in a setting which is not representative of a general dental setting. OCT appears to be a promising device, but this technology has principally been evaluated on extracted teeth in eight datasets. There is a need to investigate further the use of OCT in a general dental practice setting.

AUTHORS' CONCLUSIONS

Implications for practice

We intended that the results of this Cochrane Review be directly applicable to the general dental practitioner. Ideally, clinicians would have all diagnostic test or devices available to them and use the most appropriate according to the clinical scenario. This is not possible for most dental practices who have finite resources and existing infrastructure that would almost always feature a radiographic device to support the conventional oral examination. The question remains to clinicians whether the utilisation of a transillumination or optical coherence tomography (OCT) device provides sufficient benefit to justify the cost.

Transillumination and OCT devices appear to be valuable tools in the detection of enamel caries, with evidence suggesting that OCT devices are superior to near-infrared (NIR) transillumination and fibre-optic transillumination (FOTI)/digital fibre-optic transillumination (DIFOTI). However, OCT devices are still relatively early in their development, and further research and refinement of the devices are required to enable effective technology transfer to general dental practitioners. Another consideration is the training requirement for the implementation of OCT, it is claimed to have a short training time ([Holtzman 2010](#)), but the device is unique in its required method of interpretation. Given that all undergraduate dentists currently learn to interpret radiographs, it is crucial to understand the utility of this device if retraining of the workforce is required.

The ability to confirm that a tooth is healthy (high specificity) using a given detection method is of particular importance to the general dental practitioner to avoid the provision of unnecessary interventional treatment. This requirement is balanced against the consequences of incorrectly classifying a tooth as healthy when enamel caries exists (low sensitivity), which may lead to missed opportunities to apply enhanced preventative strategies and an increased risk of lesion progression. The results from the meta-analysis have indicated that FOTI/DIFOTI has the highest specificity and therefore could be considered as a useful tool to detect sound or healthy teeth, particularly when considered as an adjunct to a visual examination. This device also benefits from being a non-ionising alternative to radiography. Similar benefits are observed for NIR devices, with the addition of longitudinal evaluation made possible due to the inclusion of imaging capabilities in these devices. The recent inclusion of NIR technology in intraoral scanners, which are increasingly utilised by general practitioners, further elevates the importance of understanding the applicability of this modality for the detection of early caries. FOTI/DIFOTI is also one of the simplest, most inexpensive caries diagnostic aids available and is more readily implementable than OCT. It has the added benefit of being applicable for anterior and posterior teeth, whereas NIR devices are typically intended for posterior teeth. More generally, the application of this knowledge should be considered within the context of caries risk assessment in practice, which should consider various factors beyond the detection of an early lesion for the determination of the most appropriate preventive strategy for the patient. Applying best practice prevention may help to mitigate the risks of false negative findings from the application of FOTI and NIR technologies within a framework of holistic patient management.

The reproducibility of the devices was beyond the remit of this review. One useful clinical application of these devices would be to use the imaging component to record lesions at multiple time points to allow them to be monitored for severity and activity and to justify any intervention. Clinicians will always perform a visual examination but then look to adjunct methods to provide validation or confirmation of their decision. The results do appear to suggest that illumination devices will add value to the diagnostic decision process.

Implications for research

Further research is required on OCT, particularly in a clinical setting and investigating its ability to detect caries on approximal and occlusal surfaces. Greater understanding and a direct comparison of the mode of action of OCT (such as 'swept source' versus 'spectral domain') is required to develop this into a commercially viable device that can easily be deployed and interpreted by the clinician. We would question the need for any further research on FOTI or DIFOTI. It may be possible to view this as an adjunct to a visual examination to illuminate the area further and can easily be carried out using the dental mirror and light in the dental practice. Although we should highlight that the FOTI tools employed in the studies reported in this review produce a narrow beam fibre optic which is believed to offer benefits over the reflection of overhead mirror lighting. High-quality studies for NIR could still be useful as these are new devices, and their true effect on the dental examination does not currently seem to have been explored.

As highlighted by the number of studies excluded from this review for incomplete reporting of outcome data, it is of vital importance that future research studies report the data in a clear, concise form that addresses each item of the STARD checklist ([Bossuyt 2003](#); [Cohen 2015](#)). Ideally this would include a cross-tabulation of the index test and reference standard with a minimum requirement of three categories of each which could be classified as sound/caries free, early/enamel caries, advanced/dentine caries. Many studies subdivided these latter two categories into inner and outer enamel/dentine caries, and this allowed us to extract true-positive, false-positive, false-negative, and true-positive results.

Importantly, future studies should be aware of the importance of sampling participants using consecutive or random sampling. This should serve to minimise the bias which originates from the selection of teeth in which early caries is either easier or more difficult to detect. Studies should also specify the test positivity thresholds a priori rather than selecting the threshold which maximises estimates of sensitivity and specificity, ideally using manufacturer recommended thresholds or those validated in previous research studies. Studies may be conducted to determine the most accurate thresholds for a given population. We would recommend that such studies also report the manufacturer's recommended thresholds to facilitate a comparison between the two and allow for analysis in future reviews.

One possible design of an ideal diagnostic test accuracy study would be where participants are identified that required a tooth extraction, enabling the index test to be conducted in the clinical setting, and histology being possible for the reference standard once the tooth has been extracted. Future studies could look at the potential of illumination devices to be used in combination with other technologies and to make direct comparisons between their use at different points of the disease spectrum, i.e. general practice:

seemingly asymptomatic, low/high need, irregular attenders, previously diseased participants. Given the potential utility of the devices in aiding the clinician to confirm borderline cases where the clinician is uncertain of the true disease state, a study could be designed which investigates only those sites which have a degree of uncertainty.

ACKNOWLEDGEMENTS

This series of Cochrane Reviews was funded by the UK National Institute for Health Research (NIHR) Cochrane Programme Grant Scheme (Project: 16/114/23). We would like to thank Anne

Littlewood (Information Specialist, Cochrane Oral Health) for her advice on the search strategy and conducting the search of the literature, and Luisa M Fernandez Mauleffinch (Managing Editor and Copy Editor, Cochrane Oral Health) for her assistance in facilitating this review. We thank Associate Professor KR Ekstrand, J Bader (Emeritus Professor, UNC School of Dentistry, Chapel Hill North Carolina, USA), Iain Pretty, and Patrick Fee for their feedback on the protocol; Derek Richards, J Bader, Jennifer Hilgart, Peter Tugwell, and the Cochrane Diagnostic Test Accuracy Editorial Team for their feedback on the review. Also Alex Sutton and Suzanne Freeman from the NIHR Complex Review Support Unit, and Yemisi Takwoingi from the University of Birmingham for their support on this review.

REFERENCES

References to studies included in this review

Ashley 1998 {published data only}

Ashley PF, Blinkhorn AS, Davies RM. Occlusal caries diagnosis: an in vitro histological validation of the Electronic Caries Monitor (ECM) and other methods. *Journal of Dentistry* 1998;**26**:83-8.

Astvaldsdottir 2012 {published data only}

Astvaldsdottir A, Ahlund K, Holbrook WP, de Verdier B, Tranaeus S. Approximal caries detection by DIFOTI: in vitro comparison of diagnostic accuracy/efficacy with film and digital radiography. *International Journal of Dentistry* 2012;**2012**:326401.

Bussaneli 2015 {published data only}

Bussaneli DG, Restrepo M, Boldieri T, Pretel H, Mancini MW, Santos-Pinto L, et al. Assessment of a new infrared laser transillumination technology (808 nm) for the detection of occlusal caries - an in vitro study. *Lasers in Medical Science* 2015;**30**:1873-9.

Chawla 2012 {published data only}

Chawla N, Messer LB, Adams GG, Manton DJ. An in vitro comparison of detection methods for approximal carious lesions in primary molars. *Caries Research* 2012;**46**:161-9.

Holt 1989 {published data only}

Holt RD, Azevedo MR. Fibre optic transillumination and radiographs in diagnosis of approximal caries in primary teeth. *Community Dental Health* 1989;**6**:239-47.

Holtzman 2010 {published data only}

Holtzman JS, Osann K, Pharar J, Lee K, Ahn YC, Tucker T, et al. Ability of optical coherence tomography to detect caries beneath commonly used dental sealants. *Lasers in Surgery and Medicine* 2010;**42**:752-9.

Jablonski-Momeni 2017 {published data only}

Jablonski-Momeni A, Jablonski B, Lippe N. Clinical performance of the near-infrared imaging system VistaCam iX Proxi for detection of approximal enamel lesions. *BDJ Open* 2017;**3**:17012.

Laitala 2017 {published data only}

Laitala ML, Piipari L, Sampi N, Korhonen M, Pesonen P, Joensuu T, et al. Validity of digital imaging of fiber-optic transillumination in caries detection on proximal tooth surfaces. *International Journal of Dentistry* 2017;**2017**:8289636.

Lederer 2018 {published data only}

Lederer A, Kunzelmann KH, Hickel R, Litzenburger F. Transillumination and HDR imaging for proximal caries detection. *Journal of Dental Research* 2018;**97**:844-9.

Mansour 2016 {published data only}

Mansour S, Ajdaharian J, Nabelsi T, Chan G, Wilder-Smith P. Comparison of caries diagnostic modalities: a clinical study in 40 subjects. *Lasers in Surgery and Medicine* 2016;**48**:924-8.

Mialhe 2003 {published data only}

Mialhe FL, Pereira AC, Pardi V, de Castro Meneghim M. Comparison of three methods for detection of carious lesions in proximal surfaces versus direct visual examination after tooth separation. *Journal of Clinical Pediatric Dentistry* 2003;**28**:59-62.

Nakagawa 2013 {published data only}

Nakagawa H, Sadr A, Shimada Y, Tagami J, Sumi Y. Validation of swept source optical coherence tomography (SS-OCT) for the diagnosis of smooth surface caries in vitro. *Journal of Dentistry* 2013;**41**:80-9.

Nakajima 2014 {published data only}

Nakajima Y, Shimada Y, Sadr A, Wada I, Miyashin M, Takagi Y, et al. Detection of occlusal caries in primary teeth using swept source optical coherence tomography. *Journal of Biomedical Optics* 2014;**19**:16020.

NCT02657538 {published data only}

NCT02657538. Validation of near-infrared light transillumination for interproximal enamel caries detection [Validation of near-infrared light transillumination for interproximal enamel caries detection: a clinical controlled trial]. clinicaltrials.gov/ct2/show/NCT02657538 (first received 18 January 18 2016).

Obry-Musset 1988 {published data only}

Obry-Musset AM, Cahen PM, Turlot JC, Frank RM. Approximal caries diagnosis in epidemiological studies: transillumination or bitewing radiographs? *Journal de Biologie Buccale* 1988;**16**:13-7.

Park 2018 {published data only}

Park KJ, Schneider H, Ziebolz D, Krause F, Haak R. Optical coherence tomography to evaluate variance in the extent of carious lesions in depth. *Lasers in Medical Science* 2018;**33**:1573-9.

Shimada 2010 {published data only}

Shimada Y, Sadr A, Burrow MF, Tagami J, Ozawa N, Sumi Y. Validation of swept-source optical coherence tomography (SS-OCT) for the diagnosis of occlusal caries. *Journal of Dentistry* 2010;**38**:655-65.

Shimada 2014 {published data only}

Shimada Y, Nakagawa H, Sadr A, Wada I, Nakajima M, Nikaido T, et al. Non-invasive cross-sectional imaging of proximal caries using swept-source optical coherence tomography (SS-OCT) in vivo. *Journal of Biophotonics* 2014;**7**:506-13.

Sidi 1988 {published data only}

Sidi AD, Naylor MN. A comparison of bitewing radiography and interdental transillumination as adjuncts to the clinical identification of approximal caries in posterior teeth. *British Dental Journal* 1988;**164**:15-8.

Simon 2016 {published data only}

Simon JC, Lucas SA, Staninec M, Tom H, Chan KH, Darling CL, et al. Near-IR transillumination and reflectance imaging at 1300 nm and 1500 - 1700 nm for in vivo caries detection. *Lasers in Surgery and Medicine* 2016;**48**:828-36.

Simon 2016a {published data only}

Simon JC, Lucas SA, Staninec M, Tom H, Chan KH, Darling CL, et al. Near-IR transillumination and reflectance imaging at 1300 nm and 1500 - 1700 nm for in vivo caries detection. *Lasers in Surgery and Medicine* 2016;**48**:828-36.

Van Hilsen 2013 {published data only}

Van Hilsen Z, Jones RS. Comparing potential early caries assessment methods for teledentistry. *BMC Oral Health* 2013;**13**:16.

Xiao-Hua 2016 {published data only}

Xiao-Hua D, Hui Y, Xiaoli L, Yanni L, Yingying W, Xiaobin L, et al. Ex vivo assessment of the potency of optical coherence tomography for the detection of early occlusal caries. *Hua Xi Kou Qiang Yi Xue Za Zhi [West China Journal of Stomatology]* 2016;**34**:564-9.

Zain 2018 {published data only}

Zain E, Zakian CM, Chew HP. Influence of the loci of non-cavitated fissure caries on its detection with optical coherence tomography. *Journal of Dentistry* 2018;**71**:31-7.

References to studies excluded from this review
Baltacioglu 2017 {published data only}

Baltacioglu IH, Orhan K. Comparison of diagnostic methods for early interproximal caries detection with near-infrared light transillumination: an in vivo study. *BMC Oral Health* 2017;**17**:130.

Bin-Shuwaish 2008 {published data only}

Bin-Shuwaish M, Yaman P, Dennison J, Neiva G. The correlation of DIFOTI to clinical and radiographic images in Class II carious lesions. *Journal of the American Dental Association* 2008;**139**:1374-81.

Blazejewska 2016 {published data only}

Blazejewska A, Dacyna N, Niesiobedzki P, Trzaska M, Gozdowski D, Turska-Szybka A, et al. Comparison of the detection of proximal caries in children and youth using DIAGNOcam and bitewing radiovisiography. *Dental and Medical Problems* 2016;**53**:468-75.

Chesters 2002 {published data only}

Chesters RK, Pitts NB, Matulienė G, Kvedariene A, Huntington E, Bendinskaite R, et al. An abbreviated caries clinical trial design validated over 24 months. *Journal of Dental Research* 2002;**81**:637-40.

Choksi 1994 {published data only}

Choksi SK, Brady JM, Dang DH, Rao MS. Detecting approximal dental caries with transillumination: a clinical evaluation. *Journal of the American Dental Association* 1994;**125**:1098-102.

Cortes 2000 {published data only}

Cortes DF, Ekstrand KR, Elias-Boneta AR, Ellwood RP. An in vitro comparison of the ability of fibre-optic transillumination, visual inspection and radiographs to detect occlusal caries and evaluate lesion depth. *Caries Research* 2000;**34**:443-7.

Cortes 2003 {published data only}

Cortes DF, Ellwood RP, Ekstrand KR. An in vitro comparison of a combined FOTI/visual examination of occlusal caries with other caries diagnostic methods and the effect of stain on their diagnostic performance. *Caries Research* 2003;**37**:8-16.

da Silva 2011 {published data only}

da Silva RP, Assaf AV, Pereira SM, Mialhe FL, Ambrosano GM, Meneghim Mde C, et al. Validity of caries-detection methods under epidemiological setting. *American Journal of Dentistry* 2011;**24**:363-6.

Elhennawy 2018 {published data only}

Elhennawy K, Askar H, Jost-Brinkmann PG, Reda S, Al-Abdi A, Paris S, et al. In vitro performance of the DIAGNOcam for detecting proximal carious lesions adjacent to composite restorations. *Journal of Dentistry* 2018;**72**:39-43.

Gomez 2013 {published data only}

Gomez J, Zakian C, Salsone S, Pinto SCS, Taylor A, Pretty IA, et al. In vitro performance of different methods in detecting occlusal caries lesions. *Journal of Dentistry* 2013;**41**(2):180-6.

Grossman 2002 {published data only}

Grossman ES, Cleaton-Jones PE, Cortes DF, Daya NP, Parak RB, Fatti LP, et al. Accurate diagnosis of occlusal carious lesions - a stereo microscope evaluation of clinical diagnosis. *Journal of the South African Dental Association* 2002;**57**:215-20.

Hintze 1998 {published data only}

Hintze H, Wenzel A, Danielsen B, Nyvad B. Reliability of visual examination, fibre-optic transillumination, and bitewing radiography, and reproducibility of direct visual examination following tooth separation for the identification of cavitated carious lesions in contacting approximal surfaces. *Caries Research* 1998;**32**:204-9.

Holtzman 2014 {published data only}

Holtzman JS, Ballantine J, Fontana M, Wang A, Calantog A, Benavides E, et al. Assessment of early occlusal caries pre- and post-sealant application - an imaging approach. *Lasers in Surgery and Medicine* 2014;**46**:499-507.

Holtzman 2015 {published data only}

Holtzman JS, Kohanchi D, Biren-Fetz J, Fontana M, Ramchandani M, Osann K, et al. Detection and proportion of very early dental caries in independent living older adults. *Lasers in Surgery and Medicine* 2015;**47**:683-8.

Ibusuki 2015 {published data only}

Ibusuki T, Kitasako Y, Sadr A, Shimada Y, Sumi Y, Tagami J. Observation of white spot lesions using swept source optical coherence tomography (SS-OCT): in vitro and in vivo study. *Dental Materials Journal* 2015;**34**:545-52.

Kuhnisch 2016 {published data only}

Kuhnisch J, Sochtig F, Pitchika V, Laubender R, Neuhaus KW, Lussi A, et al. In vivo validation of near-infrared light transillumination for interproximal dentin caries detection. *Clinical Oral Investigations* 2016;**20**:821-9.

Lara-Capi 2017 {published data only}

Lara-Capi C, Cagetti MG, Lingstrom P, Lai G, Cocco F, Simark-Mattsson C, et al. Digital transillumination in caries detection versus radiographic and clinical methods: an in vivo study. *Dento Maxillo Facial Radiology* 2017;**46**:20160417.

Lederer 2019 {published data only}

Lederer A, Kunzelmann KH, Heck K, Hickel R, Litzenburger F. In vitro validation of near-infrared transillumination at 780 nm for the detection of caries on proximal surfaces. *Clinical Oral Investigations* 2019;**23**(11):3933-40.

Lee 2009 {published data only}

Lee D, Fried D, Darling CL. Near-IR multi-modal imaging of natural occlusal lesions. *Proceedings of SPIE - the International Society for Optical Engineering* 2009;**7162**:71620X.

Lee 2010 {published data only}

Lee C, Lee D, Darling CL, Fried D. Non-destructive assessment of the severity of occlusal caries lesions with near-infrared imaging at 1310 nm. *Journal of Biomedical Optics* 2010;**15**:047011.

Lee 2010a {published data only}

Lee C, Darling CL, Fried D. In vitro near-infrared imaging of occlusal dental caries using germanium enhanced CMOS camera. *Proceedings of SPIE - the International Society for Optical Engineering* 2010;**7549**:75490K.

Manton 2007 {published data only}

Manton DJ, Messer LB. The effect of pit and fissure sealants on the detection of occlusal caries in vitro. *European Archives of Paediatric Dentistry* 2007;**8**:43-8.

Marianska 2015 {published data only}

Marianska I, Gontarz W, Pedzisz M, Turska-Szybka A, Olczak-Kowalczyk D. Identification and stage of severity of caries on proximal surfaces of deciduous teeth. *Dental and Medical Problems* 2015;**52**:39-46.

Marinova-Takorova 2014 {published data only}

Marinova-Takorova M, Anastasova R, Panov VE. Comparative evaluation of the effectiveness of five methods for early diagnosis of occlusal caries lesions - in vitro study. *Journal of IMAB - Annual Proceeding (Scientific Papers)* 2014;**20**:533-6.

Matsuura 2018 {published data only}

Matsuura C, Shimada Y, Sadr A, Sumi Y, Tagami J. Three-dimensional diagnosis of dentin caries beneath composite restorations using swept-source optical coherence tomography. *Dental Materials Journal* 2018;**37**:642-9.

Mialhe 2009 {published data only}

Mialhe FL, Pereira AC, Meneghim Mde C, Ambrosano GM, Pardi V. The relative diagnostic yields of clinical, FOTI and radiographic examinations for the detection of approximal caries in youngsters. *Indian Journal of Dental Research* 2009;**20**:136-40.

Mitropoulos 1985 {published data only}

Mitropoulos CM. A comparison of fibre-optic transillumination with bitewing radiographs. *British Dental Journal* 1985;**159**:21-3.

Mitropoulos 1985a {published data only}

Mitropoulos CM. The use of fibre-optic transillumination in the diagnosis of posterior approximal caries in clinical trials. *Caries Research* 1985;**19**:379-84.

Ozkan 2017 {published data only}

Ozkan G, Guzel KGU. Clinical evaluation of near-infrared light transillumination in approximal dentin caries detection. *Lasers in Medical Science* 2017;**32**(6):1417-22.

Purdell-Lewis 1974 {published data only}

Purdell-Lewis DJ, Pot T. A comparison of radiographic and fibre-optic diagnoses of approximal caries lesions. *Journal of Dentistry* 1974;**2**:143-8.

Salsone 2012 {published data only}

Salsone S, Taylor A, Gomez J, Pretty I, Ellwood R, Dickinson M, et al. Histological validation of near-infrared reflectance multispectral imaging technique for caries detection and quantification. *Journal of Biomedical Optics* 2012;**17**:076009.

Sasazawa 2015 {published data only}

Sasazawa S, Kakino S, Matsuura Y. Optical-fiber-based laser-induced breakdown spectroscopy for detection of early caries. *Journal of Biomedical Optics* 2015;**20**:065002.

Schaefer 2018 {published data only}

Schaefer G, Pitchika V, Litzenburger F, Hickel R, Kühnisch J. Evaluation of occlusal caries detection and assessment by visual inspection, digital bitewing radiography and near-infrared light transillumination. *Clinical Oral Investigations* 2018;**22**(7):2431-8.

Schneiderman 1997 {published data only}

Schneiderman A, Elbaum M, Shultz T, Keem S, Greenebaum M, Driller J. Assessment of dental caries with digital imaging fiber-optic transillumination (DIFOTI): in vitro study. *Caries Research* 1997;**31**:103-10.

Simon 2014 {published data only}

Simon JC, Lucas SA, Staninec M, Tom H, Chan KH, Darling CL, et al. Transillumination and reflectance probes for in vivo near-IR imaging of dental caries. *Proceedings of SPIE - the International Society for Optical Engineering* 2014;**8929**:89290D.

Simon 2018 {published data only}

Simon JC, Curtis DA, Darling CL, Fried D. Multispectral near-infrared reflectance and transillumination imaging of occlusal carious lesions: variation in lesion contrast with lesion depth. *Proceedings of SPIE - the International Society for Optical Engineering* 2018;**10473**:1047305.

Sochtig 2014 {published data only}

Sochtig F, Hickel R, Kühnisch J. Caries detection and diagnostics with near-infrared light transillumination: clinical experiences. *Quintessence International* 2014;**45**:531-8.

Staninec 2010 {published data only}

Staninec M, Lee C, Darling CL, Fried D. In vivo near-IR imaging of approximal dental decay at 1310 nm. *Lasers in Surgery and Medicine* 2010;**42**:292-8.

Staninec 2011 {published data only}

Staninec M, Douglas SM, Darling CL, Chan K, Kang H, Lee RC, et al. Non-destructive clinical assessment of occlusal caries lesions using near-IR imaging methods. *Lasers in Surgery and Medicine* 2011;**43**:951-9.

Stephen 1987 {published data only}

Stephen KW, Russell JI, Creanor SL, Burchell CK. Comparison of fibre optic transillumination with clinical and radiographic caries diagnosis. *Community Dentistry and Oral Epidemiology* 1987;**15**:90-4.

Tetschke 2018 {published data only}

Tetschke F, Kirsten L, Golde J, Walther J, Galli R, Koch E, et al. Application of optical and spectroscopic technologies for the characterization of carious lesions in vitro. *Biomedizinische Technik [Biomedical Engineering]* 2018;**63**:595-602.

Verdonschot 1991 {published data only}

Verdonschot EH, van de Rijke JW, Brouwer W, ten Bosch JJ, Truin GJ. Optical quantitation and radiographic diagnosis of incipient approximal caries lesions. *Caries Research* 1991;**25**:359-64.

Verdonschot 1992 {published data only}

Verdonschot EH, Bronkhorst EM, Burgersdijk RC, Konig KG, Schaeken MJ, Truin GJ. Performance of some diagnostic systems in examinations for small occlusal carious lesions. *Caries Research* 1992;**26**:59-64.

Waly 1995 {published data only}

Waly NG. Evaluation of three diagnostic methods for initial proximal caries detection in primary molars. *Egyptian Dental Journal* 1995;**41**:1441-9.

Wenzel 1992 {published data only}

Wenzel A, Verdonschot EH, Truin GJ, Konig KG. Accuracy of visual inspection, fiber-optic transillumination, and various radiographic image modalities for the detection of occlusal caries in extracted non-cavitated teeth. *Journal of Dental Research* 1992;**71**:1934-7.

Wilder-Smith 2013 {published data only}

Wilder-Smith P, Holtzman J, Fontana M, Ballantine J, Calantog A, Wang A, et al. Development and validation of a caries diagnostic algorithm. *Lasers in Surgery and Medicine* 2013;**45**:274.

Yu 2017 {published data only}

Yu JL, Tang RT, Feng L, Dong YM. Digital imaging fiber optic transillumination (DIFOTI) method for determining the depth of cavity. *Beijing Da Xue Xue Bao [Journal of Peking University. Health sciences]* 2017;**49**:81-5.

Zakian 2009 {published data only}

Zakian C, Pretty I, Ellwood R. Near-infrared hyperspectral imaging of teeth for dental caries detection. *Journal of Biomedical Optics* 2009;**14**:064047.

Additional references
Anaise 1984

Anaise JZ. Measurement of dental caries experience - modification of the DMFT index. *Community Dentistry and Oral Epidemiology* 1984;**12**(1):43-6.

Angmar-Månsson 2001

Angmar-Månsson B, Ten Bosch JJ. Quantitative light-induced fluorescence (QLF): a method for assessment of incipient caries lesions. *Dento Maxillo Facial Radiology* 2001;**30**:298-307.

Backer Dirks 1951

Backer Dirks O, Van Amerongen J, Winkler KC. A reproducible method for caries evaluation. *Journal of Dental Research* 1951;**30**:346-59.

Bader 2002

Bader JD, Shugars DA, Bonito AJ. A systematic review of the performance of methods for identifying carious lesions. *Journal of Public Health Dentistry* 2002;**62**(4):201-13.

Bakhshandeh 2018

Bakhshandeh A, Floriano I, Braga MM, Thorlacius KA, Ekstrand KR. Relationship between depth of approximal caries lesions and presence of bacteria in the dentine in primary and permanent posterior teeth: a radiographic examination with microbiological evaluation. *Acta Odontologica Scandinavica* 2018;**76**(7):509-14.

Bates 2007

Bates D, Sarkar D, Bates MD, Matrix L. The lme4 package. *R Package Version* 2007;**2**(1):74.

Bloemendal 2004

Bloemendal E, De Vet HC, Bouter LM. The value of bitewing radiographs in epidemiological caries research: a systematic review of the literature. *Journal of Dentistry* 2004;**32**:255-64.

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Annals of Internal Medicine* 2003;**138**(1):W1-12.

Braga 2009

Braga MM, Oliveira LB, Bonini GAVC, Bönecker M, Mendes FM. Feasibility of the International Caries Detection and Assessment System (ICDAS-II) in epidemiological surveys and comparability with standard World Health Organization criteria. *Caries Research* 2009;**43**:245-9.

Broadbent 2008

Broadbent JM, Thomson WM, Poulton R. Trajectory patterns of dental caries experience in the permanent dentition to the fourth decade of life. *Journal of Dental Research* 2008;**87**:69-72.

Cohen 2015

Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic

accuracy studies: explanation and elaboration. *BMJ Open* 2016;**6**(11):e012799.

Davies 2001

Davies GM, Worthington HV, Clarkson JE, Thomas P, Davies RM. The use of fibre-optic transillumination in general dental practice. *British Dental Journal* 2001;**191**:145.

Deeks 2005

Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of Clinical Epidemiology* 2005;**58**(9):882-93.

Deeks 2013

Deeks JJ, Bossuyt PM, Gatsonis C, editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration, 2013. Available from srdta.cochrane.org.

de Vet 2008

de Vet HCW, Eisinga A, Riphagen II, Aertgeerts B, Pewsner D. Chapter 7: Searching for studies. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 (updated September 2008). The Cochrane Collaboration, 2008. Available from srdta.cochrane.org.

Dinnes 2016

Dinnes J, Mallett S, Hopewell S, Roderick PJ, Deeks JJ. The Moses-Littenberg meta-analytical method generates systematic differences in test accuracy compared to hierarchical meta-analytical models. *Journal of Clinical Epidemiology* 2016;**80**:77-87.

Downer 1975

Downer MC. Concurrent validity of an epidemiological diagnostic system for caries with the histological appearance of extracted teeth as validating criterion. *Caries Research* 1975;**9**:231-46.

Dye 2015

Dye BA, Thornton-Evans G, Li X, Iafolla TJ. Dental caries and sealant prevalence in children and adolescents in the United States, 2011-2012. *NCHS Data Brief* 2015;**(191)**:1-8.

Ekstrand 1997

Ekstrand KR, Ricketts DNJ, Kidd EAM. Reproducibility and accuracy of three methods for assessment of demineralization depth on the occlusal surface: an in vitro examination. *Caries Research* 1997;**31**:224-31.

Ekstrand 1998

Ekstrand KR, Ricketts DNJ, Kidd EAM, Qvist V, Schou S. Detection, diagnosing, monitoring and logical treatment of occlusal caries in relation to lesion activity and severity: an in vivo examination with histological validation. *Caries Research* 1998;**32**:247-54.

Ekstrand 2007

Ekstrand KR, Martignon S, Ricketts DJN, Qvist V. Detection and activity assessment of primary coronal caries lesions: a methodologic study. *Operative Dentistry* 2007;**32**(3):225-35.

Ellwood 1997

Ellwood RP, Davies RM, Worthington HV. Evaluation of a dental subtraction radiography system. *Journal of Periodontal Research* 1997;**32**:241-8.

Featherstone 2004

Featherstone JDB. The continuum of dental caries - evidence for a dynamic disease process. *Journal of Dental Research* 2004;**83**:39-42.

Feigin 2016

Feigin V. Global, regional, and national incidence, prevalence, and years lived with disability for 310 acute and chronic diseases and injuries, 1990-2015: a systematic analysis for the global burden of disease study 2015. *Lancet* 2016;**388**:1545-602.

Fejerskov 2015

Fejerskov O, Nyvad B, Kidd E, editor(s). Dental Caries: the Disease and its Clinical Management. 3rd edition. Wiley-Blackwell, 2015.

Freeman 2019

Freeman SC, Kerby CR, Patel A, Cooper NJ, Quinn T, Sutton AJ. Development of an interactive web-based tool to conduct and interrogate meta-analysis of diagnostic test accuracy studies: MetaDTA. *BMC Medical Research Methodology* 2019;**19**(1):81.

Fried 2019

Fried D. Optical coherence tomography for imaging dental caries. In: Ferreira Zandona A, Longbottom C, editors(s). Detection and Assessment of Dental Caries. Springer, Cham, 2019:199-208.

Fyffe 2000

Fyffe HE, Deery C, Nugent ZJ, Nuttall NM, Pitts NB. Effect of diagnostic threshold on the validity and reliability of epidemiological caries diagnosis using the Dundee Selectable Threshold Method for caries diagnosis (DSTM). *Community Dentistry and Oral Epidemiology* 2000;**28**:42-51.

Hall-Scullin 2017

Hall-Scullin E, Whitehead H, Milsom K, Tickle M, Su T-L, Walsh T. Longitudinal study of caries development from childhood to adolescence. *Journal of Dental Research* 2017;**96**:762-7.

Hogan 2019

Hogan R, Pretty IA, Ellwood RP. Fibre-optic transillumination: FOTI. In: Ferreira Zandona A, Longbottom C, editors(s). Detection and Assessment of Dental Caries. Springer, Cham, 2019:139-50.

Horner 2009

Horner K, Islam M, Flygare L, Tsiklakis K, Whaites E. Basic principles for use of dental cone beam computed tomography: consensus guidelines of the European Academy of Dental

and Maxillofacial Radiology. *Dento Maxillo Facial Radiology* 2009;**38**:187-95.

Hse 2015

Hse K. The challenge of oral disease - a call for global action. In: The Oral Health Atlas. 2nd edition. Geneva: FDI World Dental Federation, 2015:8-11.

Ismail 2004

Ismail AI. Visual and visuo-tactile detection of dental caries. *Journal of Dental Research* 2004;**83**:56-66.

Ismail 2007

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 Dentistry and Oral Epidemiology* 2007;**35**:170-8.

Ismail 2013

Ismail AI, Tellez M, Pitts NB, Ekstrand KR, Ricketts D, Longbottom C, et al. Caries management pathways preserve dental tissues and promote oral health. *Community Dentistry and Oral Epidemiology* 2013;**41**(1):e12-40.

Ismail 2015

Ismail AI, Pitts NB, Tellez M. The International Caries Classification and Management System (ICCMS™) an example of a caries management pathway. *BMC Oral Health* 2015;**15** Suppl 1:S9. [DOI: [10.1186/1472-6831-15-S1-S9](https://doi.org/10.1186/1472-6831-15-S1-S9)]

Jones 2017

Jones CM, Davies GM, Monaghan N, Morgan MZ, Neville JS, Pitts NB. The caries experience of 5 year-old children in Scotland in 2013-2014, and in England and Wales in 2014-2015. Reports of cross-sectional dental surveys using BASCD criteria. *Community Dental Health* 2017;**34**:157-62.

Kassebaum 2015

Kassebaum NJ, Bernabé E, Dahiya M, Bhandari B, Murray CJL, Marcenes W. Global burden of untreated caries: a systematic review and metaregression. *Journal of Dental Research* 2015;**94**:650-8.

Kidd 2004

Kidd EAM, Fejerskov O. What constitutes dental caries? Histopathology of carious enamel and dentin related to the action of cariogenic biofilms. *Journal of Dental Research* 2004;**83**:35-8.

Kidd 2016

Kidd EAM, Fejerskov O. Essentials of Dental Caries. 4th edition. Oxford University Press, 2016.

Kühnisch 2019

Kühnisch J. Near-infrared light transillumination. In: Ferreira Zandona A, Longbottom C, editors(s). Detection and Assessment of Dental Caries. Springer, Cham, 2019:151-8.

Leeflang 2008

Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine* 2008;**149**:889-97.

Listl 2015

Listl S, Galloway J, Mossey PA, Marcenes W. Global economic impact of dental diseases. *Journal of Dental Research* 2015;**94**:1355-61.

Macaskill 2010

Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C, editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010. Available from srdta.cochrane.org.

Macey 2018

Macey R, Walsh T, Riley P, Glenny AM, Worthington H, Clarkson J, et al. Tests to detect and inform the diagnosis of caries. *Cochrane Database of Systematic Reviews* 2018, Issue 12. Art. No: CD013215. [DOI: [10.1002/14651858.CD013215](https://doi.org/10.1002/14651858.CD013215)]

Macey 2020

Macey R, Walsh T, Riley P, Glenny AM, Worthington HV, Fee PA, et al. Fluorescence devices for the detection of dental caries. *Cochrane Database of Systematic Reviews* 2020, Issue 12. Art. No: CD013811. [DOI: [10.1002/14651858.CD013811](https://doi.org/10.1002/14651858.CD013811)]

Martignon 2019

Martignon S, Pitts NB, Goffin G, Mazevet M, Douglas GVA, Newton JT, et al. CariesCare practice guide: consensus on evidence into practice. *British Dental Journal* 2019;**227**(5):353-62.

Matos 2011

Matos R, Novaes TF, Braga MM, Siqueira WL, Duarte DA, Mendes FM. Clinical performance of two fluorescence-based methods in detecting occlusal caries lesions in primary teeth. *Caries Research* 2011;**45**:294-302.

McInnes 2018

McInnes MDF, Moher D, Thoms BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA* 2018;**319**:388-96.

Milsom 2003

Milsom KM, Tickle M, Humphris GM, Blinkhorn AS. The relationship between anxiety and dental treatment experience in 5-year-old children. *British Dental Journal* 2003;**194**:503.

Nyvad 1997

Nyvad B, Fejerskov O. Assessing the stage of caries lesion activity on the basis of clinical and microbiological examination. *Community Dentistry and Oral Epidemiology* 1997;**25**:69-75.

Partlett 2016

Partlett C, Takwoingi Y. Meta-analysis of test accuracy studies in R: a summary of user-written programs and step-by-step guide to using glmer Version 1.0. srdata.cochrane.org/ 2016.

Petersen 2005

Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. *Bulletin of the World Health Organization* 2005;**83**:661-9.

Pitts 1988

Pitts NB, Fyffe HE. The effect of varying diagnostic thresholds upon clinical caries data for a low prevalence group. *Journal of Dental Research* 1988;**67**:592-6.

Pitts 2001

Pitts NB. Clinical diagnosis of dental caries: a European perspective. *Journal of Dental Education* 2001;**65**:972-8.

Pitts 2009

Pitts N. Detection, Assessment, Diagnosis and Monitoring of Caries. Switzerland: Karger Basel, 2009.

Pitts 2017

Pitts NB, Zero DT, Marsh PD, Ekstrand K, Weintraub JA, Ramos-Gomez F, et al. Dental caries. *Nature Reviews Disease Primers* 2017;**3**:17030.

Pretty 2006

Pretty IA. Caries detection and diagnosis: novel technologies. *Journal of Dentistry* 2006;**34**:727-39.

Public Health England 2014

Public Health England. Oral health survey of five-year-old. Dental public health epidemiology programme. www.nwph.net/dentalhealth/14_15_5yearold/Protocol_2014_15_5%20yr%20olds%20v2.pdf (accessed 15 May 2018).

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**:982-90.

Review Manager 2020 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.4. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2020.

Rutter 2001

Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in Medicine* 2001;**20**(19):2865-84.

Schünemann 2020

Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *Journal of Clinical Epidemiology* 2020;**122**:129-41.

Schünemann 2020a

Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2020;**122**:142-52.

Shoaib 2009

Shoaib L, Deery CM, Ricketts DNJ, Nugent ZJ. Validity and reproducibility of ICDAS II in primary teeth. *Caries Research* 2009;**43**:442-8.

Steele 2011

Steele J, O'Sullivan I. Executive summary: Adult Dental Health Survey 2009. files.digital.nhs.uk/publicationimport/pub01xxx/pub01086/adul-dent-heal-surv-summ-them-exec-2009-rep2.pdf 2011.

Takwoingi 2013

Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. *Annals of Internal Medicine* 2013;**158**(7):544-54.

Tam 2001

Tam LE, McComb D. Diagnosis of occlusal caries: Part II. Recent diagnostic technologies. *Journal of the Canadian Dental Association* 2001;**67**:459-64.

Thomson 2000

Thomson WM, Locker D, Poulton R. Incidence of dental anxiety in young adults in relation to dental treatment experience. *Community Dentistry and Oral Epidemiology* 2000;**28**:289-94.

Vernazza 2016

Vernazza CR, Rolland SL, Chadwick B, Pitts N. Caries experience, the caries burden and associated factors in children in England, Wales and Northern Ireland 2013. *British Dental Journal* 2016;**221**(6):315.

Wenzel 1995

Wenzel A, Gröndahl HG. Direct digital radiography in the dental office. *International Dental Journal* 1995;**45**:27-34.

Wenzel 2000

Wenzel A, Anthonisen PN, Juul MB. Reproducibility in the assessment of caries lesion behaviour: a comparison between conventional film and subtraction radiography. *Caries Research* 2000;**34**:214-8.

Wenzel 2004

Wenzel A. Bitewing and digital bitewing radiography for detection of caries lesions. *Journal of Dental Research* 2004;**83**:72-5.

Wenzel 2006

Wenzel A. A review of dentists' use of digital radiography and caries diagnosis with digital systems. *Dento Maxillo Facial Radiology* 2006;**35**:307-14.

Whaites 2013

Whaites E, Drage N. Essentials of Dental Radiography and Radiology. 5th edition. Churchill Livingstone Elsevier, 2013.

White 2012

White DA, Tsakos G, Pitts NB, Fuller E, Douglas GVA, Murray JJ, et al. Adult Dental Health Survey 2009: common oral health conditions and their impact on the population. *British Dental Journal* 2012;**213**(11):567.

Whiting 2011

Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality

assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**:529-36.

Wilson 1968

Wilson JMG, Jungner G, World Health Organization. Principles and Practice of Screening for Disease. Geneva: World Health Organization, 1968.

World Health Organization 2017

World Health Organization. Sugars and dental caries. WHO Technical Information Note. apps.who.int/iris/bitstream/handle/10665/259413/WHO-NMH-NHD-17.12-eng.pdf?sequence=1 (accessed 15 May 2018).

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Ashley 1998

Study characteristics

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: no cavitation and early lesions</p> <p>Teeth: permanent premolars and molars</p> <p>Sealants: excluded</p> <p>Restorations: excluded</p> <p>Surface: occlusal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: UK</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 103 teeth, 103 sites</p> <p>Prevalence: enamel 0.60, dentine 0.36</p>
Index tests	<p>Category of test: FOTI</p> <p>"tooth dried for 20 s with compressed air"</p> <p>Sequence of test(s): index tests (ECM, FOTI, radiograph) performed prior to reference standard</p> <p>Examiner training and calibration: 1 examiner: tests in a random order</p> <p>Threshold applied: sound, enamel, dentine</p> <p>Device specifics: "using a fibre-optic transillumination probe with a tip diameter of 0.5 mm (Eurotek Optical Fibre Ltd, Blackpool, UK)"</p>

Ashley 1998 (Continued)

Target condition and reference standard(s)	Category: histology		
	Sequence of index test and reference standard: index test then reference standard		
	Training of examiner: not reported		
	Blinding to index test: unclear		
	Multiple tests: no		
	Site selection: sectioned teeth		
	Target condition: sound, enamel, dentine		
Flow and timing	Participants with index test but no reference standard: 0		
	Participants with reference standard but no index test: 0		
	Time interval between tests: minimal		
	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Ashley 1998 (Continued)

Are there concerns that the index test, its conduct, or interpretation differ from the review question?

Low concern

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?

Yes

Were the reference standard results interpreted without knowledge of the results of the index tests?

Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias?

Low risk

Are there concerns that the target condition as defined by the reference standard does not match the question?

Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

Yes

Could the patient flow have introduced bias?

Low risk

Astvaldsdottir 2012

Study characteristics

Patient Sampling

Method of sampling: selected

Included conditions: no cavitation and early lesions; "The approximal surfaces of the selected teeth presented a range of conditions, from sound to non-cavitated and cavitated caries lesions"

Teeth: permanent premolars

Sealants: not reported

Restorations: not reported

Surface: approximal

Patient characteristics and setting

Age: not reported

Sex: not reported

Ethnicity: not reported

Country: Sweden

Setting: extracted teeth

Number of participants/teeth/sites: 56 teeth, 112 sites

Astvaldsdottir 2012 (Continued)

	Prevalence: enamel 0.64, dentine 0.15
Index tests	<p>Category of test: DIFOTI images</p> <p>Sequence of test(s): index tests (radiographs and transillumination) performed prior to reference standard</p> <p>Examiner training and calibration: 8 experienced examiners: "all observers underwent a 15-minute training session to become familiar with the technique"</p> <p>Threshold applied: sound outer/inner enamel, outer/inner dentine</p> <p>Device specifics: "All DIFOTI images were captured by a trained operator under standardized darkroom conditions. The DIFOTI instrument (Electro-Optical Sciences Inc, NY) was used as recommended by the manufacturer"</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: index test then reference standard</p> <p>Training of examiner: not reported</p> <p>Blinding to index test: unclear</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: sound, enamel, dentine</p>
Flow and timing	<p>Participants with index test but no reference standard: 15 surfaces out of 112 damaged during preparation of reference standard</p> <p>Participants with reference standard but no index test: 0</p> <p>Time interval between tests: minimal</p> <p>Participants receiving both tests but excluded from results: 0</p>
Comparative	
Notes	Results taken for examiner 1

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		

Astvaldsdottir 2012 (Continued)

Could the selection of patients have introduced bias?	High risk
Are there concerns that the included patients and setting do not match the review question?	High
DOMAIN 2: Index Test (All tests)	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear
Could the conduct or interpretation of the index test have introduced bias?	Unclear risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	Low risk

Bussanelli 2015

Study characteristics	
Patient Sampling	Method of sampling: selected

Bussanelli 2015 (Continued)

	<p>Included conditions: sound and early lesions; "One hundred and two permanent healthy or decayed human teeth (molars and premolars) were obtained and cleaned"</p> <p>Teeth: permanent premolars and molars</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: occlusal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Brazil</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 102 teeth</p> <p>Prevalence: enamel 0.70, dentine 0.19</p>
Index tests	<p>Category of test: near-infrared laser transillumination</p> <p>Sequence of test(s): index tests (radiograph, near-infrared then DIAGN-Odent pen and QLF) prior to reference standard; each index test separated by 1 week</p> <p>Examiner training and calibration: experienced, blinding of tests due to time gap between assessments of each index test</p> <p>Threshold applied: "measurements were used to calculate the contrast intensities"</p> <p>Device specifics: "A prototype manufactured by DMC Equipamentos (São Carlos, Brazil) was used"</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: index test then reference standard</p> <p>Training of examiner: not reported</p> <p>Blinding to index test: unclear</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: healthy, enamel, lesion at the dentine-enamel junction or dentinal</p>
Flow and timing	<p>Participants with index test but no reference standard: 0</p> <p>Participants with reference standard but no index test: 0</p> <p>Time interval between tests: minimal</p>

Bussanelli 2015 (Continued)

Participants receiving both tests but excluded from results: 8 teeth excluded from results as near-infrared device failed to return a result, therefore excluded from all tests

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Bussanelli 2015 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	High risk

Chawla 2012

Study characteristics

Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported Restorations: not reported Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Australia Setting: extracted teeth Number of participants/teeth/sites: 135 sites (140 originally with 5 damaged) Prevalence: enamel 0.61, dentine 0.24
Index tests	Category of test: 2 transillumination lights (FOTI) (SDI and NSK, Radii Plus and Ti-Max P200 Phatelus Probe) Sequence of test(s): index tests (visual, radiograph, transillumination then DIAGNOdent then DIAGNOdent pen) prior to reference standard Examiner training and calibration: calibration and training completed; attempt made to blind examiners to each index test Threshold applied: sound, inner/outer enamel, inner/outer dentine Device specifics: "Teeth were examined in ambient light without room darkening. Examination was performed from the buccal, occlusal and lingual aspects"
Target condition and reference standard(s)	Category: histology

Chawla 2012 (Continued)

	Sequence of index test and reference standard: index test then reference standard
	Training of examiner: not reported
	Blinding to index test: yes
	Multiple tests: no
	Site selection: sectioned teeth
	Target condition: healthy, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 5
	Participants with reference standard but no index test: 0
	Time interval between tests: minimal
	Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data used from SDI device as most comparable to other devices, figure 1 used for data extraction

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Chawla 2012 (Continued)

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Holt 1989

Study characteristics

Patient Sampling	Method of sampling: not reported Included conditions: not reported; "bitewing radiographs were indicated" Teeth: primary canines and molars Sealants: not reported Restorations: not reported Surface: approximal
Patient characteristics and setting	Age: 5 to 10 years Sex: not reported Ethnicity: not reported Country: UK Setting: community dental clinic Number of participants/teeth/sites: 54 participants, 781 surfaces Prevalence: enamel 0.08
Index tests	Category of test: FOTI Sequence of test(s): index tests prior to reference standard

Transillumination and optical coherence tomography for the detection and diagnosis of enamel caries (Review)

Holt 1989 (Continued)

Not 100% (continued)

	Examiner training and calibration: not reported - examiners viewed independently		
	Threshold applied: sound, any caries		
	Device specifics: "150H universal light source, housing a 150 watt tungsten bulb, a 2 mm diameter fibre optic cable and a fibre optic probe tip of 0.5 mm diameter"		
Target condition and reference standard(s)	Category: radiographs Sequence of index test and reference standard: reference standard completed prior to transillumination Training of examiner: not reported Blinding to index test: yes - completed after index test Site selection: all approximal surfaces Target condition: sound, enamel, or dentine		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 209 deemed unreadable by radiograph		
Comparative			
Notes	Examiner 1 results used for data analysis		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		

Holt 1989 (Continued)

If multiple tests were applied were different examiners used for each (in vivo)?

Could the conduct or interpretation of the index test have introduced bias?

Low risk

Are there concerns that the index test, its conduct, or interpretation differ from the review question?

Low concern

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? No

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Could the reference standard, its conduct, or its interpretation have introduced bias?

High risk

Are there concerns that the target condition as defined by the reference standard does not match the question?

High

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

Could the patient flow have introduced bias?

High risk

Holtzman 2010
Study characteristics

Patient Sampling	Method of sampling: selected
	Included conditions: sound, early/enamel lesions, no cavitation
	Teeth: not reported
	Sealants: results taken for no sealant (also reports results after sealing all teeth)
	Surface: not reported
Patient characteristics and setting	Age: not reported
	Sex: not reported
	Ethnicity: not reported
	Country: USA
	Setting: extracted teeth

Holtzman 2010 (Continued)

	<p>Number of participants/teeth/sites: 40 teeth, 40 sites</p> <p>Prevalence: enamel 0.65, dentine (no data)</p>
Index tests	<p>Category of test: spectral domain - OCT</p> <p>Sequence of test(s): 1) visual, 2) OCT, 3) radiography, 4) histology (examiners blinded to different tests)</p> <p>Examiner training and calibration: 2 pre-standardised scorers separate from the initial selection of the tooth samples (dentists, each with > 10 years clinical experience)</p> <p>Threshold applied: healthy/not healthy</p> <p>Device specifics: "Five hundred twelve sequential 2D-OCT images of each tooth at a wavelength of 1350 nm were taken with SD-OCT. The SD-OCT system had a stationary reference arm with the tooth sample set on a stage for imaging"</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: reference standard last</p> <p>Training of examiner: not reported</p> <p>Blinding to index test: yes (states "blinded" and "not aware of the previous diagnostic test results")</p> <p>Multiple tests: no</p> <p>Site selection: serial photomicrographs (10x) were taken of the region of interest of each section</p> <p>Target condition: presence or absence of pathology (demineralisation (very early stages in the caries process which may be reversible)/caries (more progressed non-reversible lesions)) extending into enamel or dentine</p>
Flow and timing	<p>Participants with index test but no reference standard: 0</p> <p>Participants with reference standard but no index test: 0</p> <p>Time interval between tests: minimal</p> <p>Participants receiving both tests but excluded from results: 0</p>
Comparative	
Notes	<p>Results were reported pre- and post-sealant application but we have used the pre-sealant data in this review</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		

Holtzman 2010 (Continued)

Could the selection of patients have introduced bias?	High risk
Are there concerns that the included patients and setting do not match the review question?	High
DOMAIN 2: Index Test (All tests)	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
If multiple tests were applied were different examiners used for each (in vivo)?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Jablonski-Momeni 2017
Study characteristics

Patient Sampling	Method of sampling: "selected sequentially from regular dental screenings"
------------------	--

Jablonski-Momeni 2017 (Continued)

	<p>Included conditions: sound or early enamel lesions; "...criteria for inclusion were persons with an indication justifying dental radiographs or for whom radiographs were already available that were not older than 3 months"; "Approximal surfaces with IC-DAS codes higher than 2, or enamel changes not due to caries, were not included in the study"</p> <p>Teeth: permanent premolars and molars</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: mean 29.5 years (range 18 to 45 years)</p> <p>Sex: 9 male, 9 female</p> <p>Ethnicity: not reported</p> <p>Country: Germany</p> <p>Setting: routine examination in dental clinic</p> <p>Number of participants/teeth/sites: 18 participants, 161 teeth, 193 sites</p> <p>Prevalence: enamel 0.62, dentine not reported (according to visual assessment results)</p>
Index tests	<p>Category of test: NIR - VistaCam</p> <p>Sequence of test(s): visual, radiograph, NIR all assessed by same examiners but at different times; "NIR images were classified independently from the clinical and radiographical examinations"</p> <p>Examiner training and calibration: 3 examiners arrived at consensus decision, all experienced with ICDAS and radiographs</p> <p>Threshold applied: "0 = no signs of changes in enamel; NIR 1 = wide bright strip or wedge-shaped structures within the dark translucent enamel. The lesion may reach the enamel–dentine junction; NIR 2 = wide bright strip or wedge-shaped structures, which seem to cross the enamel–dentine junction"</p> <p>Device specifics: "The optical system was placed on the row of teeth above the approximal area and images were taken according to the manufacturer's instructions using the DBSWIN software"</p>
Target condition and reference standard(s)	<p>Category: visual - "teeth were separated prior to the visual examination using a dental wedge, but without additional separation of the teeth using orthodontic separation elastics"</p> <p>Sequence of index test and reference standard: NIR not likely to effect results of reference standard</p> <p>Training of examiner: same examiners as index test, but tests interpreted independently</p> <p>Blinding to index test: performed prior to NIR</p> <p>Multiple tests: visual and radiograph performed before NIR, same examiners for all, but visual assessed before radiograph</p> <p>Site selection: all surfaces</p> <p>Target condition: sound, superficial enamel demineralisation, enamel breakdown, dentine caries</p>

Jablonski-Momeni 2017 (Continued)

Flow and timing

Participants with index test but no reference standard: 0

Participants with reference standard but no index test: 8

Time interval between tests: minimal

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		

Jablonski-Momeni 2017 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	Low risk

Laitala 2017

Study characteristics	
Patient Sampling	Method of sampling: consecutive; "All 18–30-year-old university students" Included conditions: unclear Teeth: permanent premolars and molars Sealants: not reported Restorations: not reported Surface: approximal
Patient characteristics and setting	Age: 18 to 30 years Sex: not reported Ethnicity: not reported Country: Finland Setting: routine examination in dental clinic Number of participants/teeth/sites: 137 participants, of these 91 resulted in data, 1162 teeth, 2103 surfaces Prevalence: enamel 0.20, dentine 0.06 (according to radiograph results)
Index tests	Category of test: DIFOTI method

Laitala 2017 (Continued)

Sequence of test(s): visual, radiograph, DIFOTI all recorded clinically and viewed after the event by different clinicians but uncertain whether same examiner performed each test or whether tests could be influenced by others

Examiner training and calibration: trained by representatives of device manufacturer

Threshold applied: "Score 0 = no caries; Score 1 = caries lesion in outer surface of the enamel; Score 2 = caries lesion extending into the inner enamel or dento-enamel junction; Score 3 = dentinal caries lesion in the external half of dentin; Score 4 = deep dentinal caries lesion extending into the dentin half near the pulp"

Device specifics: not reported

Target condition and reference standard(s)

Category: radiograph

Sequence of index test and reference standard: visual then radiograph performed and interpreted separately to DIFOTI (index test)

Training of examiner: experienced examiner in consensus

Blinding to index test: yes

Multiple tests: visual and radiograph performed before to index test

Site selection: all surfaces

Target condition: sound, initial and manifested

Flow and timing

Participants with index test but no reference standard: 0

Participants with reference standard but no index test: possibly there is a difference in numbers that received DIFOTI and other tests, 2083 versus 2103, but unclear how this was dealt with

Time interval between tests: minimal

Participants receiving both tests but excluded from results: 0

Comparative
Notes
Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	

Laitala 2017 (Continued)

Are there concerns that the included patients and setting do not match the review question? Low concern

DOMAIN 2: Index Test (All tests)

Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

If a threshold was used, was it pre-specified? Yes

If multiple tests were applied were different examiners used for each (in vivo)? Yes

Could the conduct or interpretation of the index test have introduced bias? Unclear risk

Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? No

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Are there concerns that the target condition as defined by the reference standard does not match the question? High

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Unclear

Were all patients included in the analysis? Unclear

Could the patient flow have introduced bias? Low risk

Lederer 2018
Study characteristics

Patient Sampling	Method of sampling: selected
	Included conditions: sound or early enamel lesions; "either an intact, sound interproximal surface or non-cavitated proximal caries lesions"

Lederer 2018 (Continued)

	Teeth: permanent Sealants: not reported Restorations: excluded Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Germany Setting: extracted teeth Number of participants/teeth/sites: 53 teeth, 106 teeth Prevalence: enamel 0.36, dentine 0.16
Index tests	Category of test: 2 tests: NIR light - DIAGNOcam, also near-infrared light combined with high-dynamic-range imaging (NIR-HDRI) Sequence of test(s): NIR and NIR-HDRI, prior to reference standard Examiner training and calibration: "2 experienced examiners were given a practical and theoretical training session", examiner assessments completed independently and repeated after 2 weeks Threshold applied: no decay, enamel, enamel with single point of contact to dentino-enamel junction, enamel with extensive contact to dentino-enamel junction, dentine Device specifics: <ul style="list-style-type: none"> • DIAGNOcam - "Instead of the clinical tip of the DC, the laser optical unit of the in vitro model was used for transillumination to obtain comparable images. The DC could be pivoted over the tooth, and images obtained with the KaVo KiD (KaVo integrated Desktop) software" • NIR-HDRI - "Images were obtained with an above-mounted NIR camera manufactured by IDS (UI-1241LE-NIR-GL, Imaging Development Systems GmbH) together with a Tamron 50-mm lens" Note - applicability is "high concern" as techniques used in paper are specifically designed scenarios established in an in vitro setting which are not currently available for a clinical setting
Target condition and reference standard(s)	Category: microCT Sequence of index test and reference standard: reference standard followed index test assessments Training of examiner: unclear; "Measurements were taken 3 times by one examiner, and the mean values calculated" Blinding to index test: unclear Multiple tests: no Site selection: extracted tooth microCT scan Target condition: radiolucency in the outer, or the inner half of the enamel, and radiolucency in the dentine

Lederer 2018 (Continued)

Flow and timing

Participants with index test but no reference standard: 0

Participants with reference standard but no index test: 0

Time interval between tests: minimal

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Lederer 2018 (Continued)

Could the reference standard, its conduct, or its interpretation have introduced bias?

Unclear risk

Are there concerns that the target condition as defined by the reference standard does not match the question?

Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Could the patient flow have introduced bias?

Low risk

Mansour 2016
Study characteristics

Patient Sampling	<p>Method of sampling: unclear</p> <p>Included conditions: no cavitation and early cavitation; "Subjects with open cavities extending into dentin were excluded"</p> <p>Teeth: permanent molars</p> <p>Sealants: not reported</p> <p>Surface: "all coronal areas of the teeth considered to be at high risk of caries: occlusal and approximal, white or brown spot lesions, non-cavitated and cavitated potential lesions, fissures, and adjacent to restorations"</p>
Patient characteristics and setting	<p>Age: 19 to 52 years, mean 34</p> <p>Sex: 16 male, 24 female</p> <p>Ethnicity: not reported</p> <p>Country: USA</p> <p>Setting: dental clinic</p> <p>Number of participants/teeth/sites: 40 participants, 932 teeth (426 untreated teeth used in this sample)</p> <p>Prevalence: untreated teeth: enamel 0.12; previously treated: enamel 0.14</p>
Index tests	<p>Category of test: prototype SS-OCT system</p> <p>Sequence of test(s): index tests (laser fluorescence also completed and potentially interpreted by same examiner) performed after reference standard</p> <p>Examiner training and calibration: 90-minute training session - "Two blinded, pre-standardized examiners reviewed radiographic and OCT images independently"</p>

Mansour 2016 (Continued)

Threshold applied: examiner assessed; "OCT images independently and assigned caries status as either healthy or carious"

Device specifics: "used in this study utilizes a broadband light source with an output power of 4 mW at the center wavelength of $\lambda = 1310$ nm and bandwidth of $\Delta\lambda = 58$ nm"

Note: applicability concerns due to device being a prototype model which is not commercially available

Target condition and reference standard(s)	<p>Category: visual and radiograph; "detailed dental examination by one experienced clinician using loupes (2.5 magnification), and radiographs according to standard clinical practice"</p> <p>Sequence of index test and reference standard: index test followed reference standard</p> <p>Training of examiner: 90-minute training session</p> <p>Blinding to index test: yes, reference standard before index tests</p> <p>Multiple tests: yes</p> <p>Site selection: clinical examination</p> <p>Target condition: "Teeth were considered carious if there were white or brown spot lesions on the tooth not consistent with the clinical appearance of sound enamel"; 'healthy' being scored if both observers scored 'healthy', and 'not-healthy' scored if 1 or both observers scored 'not-healthy'</p>
--	---

Flow and timing	<p>Participants with index test but no reference standard: 0</p> <p>Participants with reference standard but no index test: 0</p> <p>Time interval between tests: minimal</p> <p>Participants receiving both tests but excluded from results: 0</p>
-----------------	---

Comparative

Notes	Untreated teeth used in the data extraction for analysis
-------	--

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern

Mansour 2016 (Continued)

DOMAIN 2: Index Test (All tests)

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

If a threshold was used, was it pre-specified? Yes

If multiple tests were applied were different examiners used for each (in vivo)? No

Could the conduct or interpretation of the index test have introduced bias? High risk

Are there concerns that the index test, its conduct, or interpretation differ from the review question? High

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? No

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Could the patient flow have introduced bias? Low risk

Mialhe 2003
Study characteristics

Patient Sampling Method of sampling: unclear
Included conditions: unclear; "low caries prevalence (DMFT = 1.4)"

Transillumination and optical coherence tomography for the detection and diagnosis of enamel caries (Review)

Mialhe 2003 (Continued)

	Teeth: permanent premolars and molars Sealants: not reported Restorations: excluded Surface: approximal Note: "surfaces considered sound by all methods were excluded from the study sample"
Patient characteristics and setting	Age: 13 to 15 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: school-based recruitment Number of participants/teeth/sites: 70 participants, 199 teeth (in the final sample) Prevalence: enamel (no-cavitation) 0.86; dentine (cavitation) 0.14
Index tests	Category of test: FOTI Sequence of test(s): index tests (visual, radiograph then FOTI) performed prior to reference standard Examiner training and calibration: 1 examiner for each test (visual, radiograph and FOTI) Threshold applied: no shadow, restricted to enamel, reaching dentine Device specifics: "FOTI (Fiber-Lite® PL 800 series (Dolan-Jenner Industries, Lawrence, MA, USA)) and a 0.5-mm-diameter probe"
Target condition and reference standard(s)	Category: visual after separation; "an orthodontic rubber ring (G & H Wire Company, Greenwood, IN) was placed with dental floss tied around the contact point of the teeth surfaces in which carious lesions had been detected. The rings were removed 24 h later and the surfaces were cleaned with dental floss and dried before examination" Sequence of index test and reference standard: reference standard followed index test Training of examiner: 3 examiners jointly decided Blinding to index test: unlikely Multiple tests: no Site selection: visual assessment of surface Target condition: sound no cavitation, cavitation
Flow and timing	Participants with index test but no reference standard: 1481 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0

Mialhe 2003 (Continued)

Note: 1481 of 1680 were excluded from analysis due to being sound

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Mialhe 2003 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	High risk

Nakagawa 2013

Study characteristics

Patient Sampling	<p>Method of sampling: "arbitrarily chosen for investigation"</p> <p>Included conditions: no cavitation and early lesions; "visible localized enamel surface discoloration (white, brown/black) or cavitated carious lesions involving an axial smooth enamel surface", therefore an attempt to have no sound surfaces</p> <p>Teeth: not reported</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: smooth</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Japan</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 93 teeth, 127 sites</p> <p>Prevalence: enamel 0.75, dentine 0.31</p>
Index tests	<p>Category of test: SS-OCT</p> <p>Sequence of test(s): index tests performed (visual, SS-OCT) prior to reference standard; "Visual examination and SS-OCT evaluations were performed in separate sessions and after shuffling the order of appearance for each case to ensure there was no interference from the previous observations"</p> <p>Examiner training and calibration: 4 dentists, 2 with 9 years experience trained the other 2, with calibration</p> <p>Threshold applied: sound, enamel demineralisation without surface breakdown, enamel breakdown due to caries, dentine caries</p> <p>Device specifics: SS-OCT (OCT-2000, Santec, Komaki, Japan); "The lateral resolution of 17 mm is determined by the objective lens of the probe. A 2000 1024</p>

Nakagawa 2013 (Continued)

	pixel image is obtained in the real-time and processed in a few hundred milliseconds"
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: index test then reference standard</p> <p>Training of examiner: "Two examiners with sufficient experience in histopathological study of caries"</p> <p>Blinding to index test: unclear</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth, through locations used for index test</p> <p>Target condition: sound, superficial enamel demineralisation, enamel breakdown, dentine caries</p>
Flow and timing	<p>Participants with index test but no reference standard: 0</p> <p>Participants with reference standard but no index test: 0</p> <p>Time interval between tests: minimal</p> <p>Participants receiving both tests but excluded from results: 0</p>

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		

Nakagawa 2013 (Continued)

Could the conduct or interpretation of the index test have introduced bias?	Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Nakajima 2014

Study characteristics

Patient Sampling	<p>Method of sampling: "Twenty-six primary teeth were selected"</p> <p>Included conditions: sound and early lesions; "38 investigation sites of occlusal fissures (non-cavitated and cavitated) were selected" - level of cavitation uncertain</p> <p>Teeth: primary molars; "Extracted human primary molar teeth with/without occlusal caries"</p> <p>Sealants: not reported Restorations: not reported</p> <p>Surface: occlusal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p>

Nakajima 2014 (Continued)

Country: Japan

Setting: dental hospital - extracted teeth

Number of participants/teeth/sites: 26 teeth, 38 sites

Prevalence: enamel 0.74, dentine 0.32

Index tests

Category of test: SS-OCT

Sequence of test(s): index tests performed (visual, OCT) prior to reference standard; "Visual examination and SS-OCT evaluations were performed in separate sessions and after shuffling the order of appearance for each case to ensure there was no interference from the previous observations"

Examiner training and calibration: 6 dentists, with 1 hour training session

Threshold applied: sound, superficial enamel demineralisation, enamel breakdown, dentine caries

Device specifics: "SS-OCT system (Prototype 2; Panasonic Healthcare Co Ltd, Ehime, Japan)"; "the center wavelength is 1330 nm (bandwidth 110 nm) with a 30-kHz sweep rate. The system is equipped with a hand-held probe with power of < 10.0 mW"

Target condition and reference standard(s)

Category: histology

Sequence of index test and reference standard: index test then reference standard

Training of examiner: "Two examiners with sufficient experience in histopathological study of caries"

Blinding to index test: unclear

Multiple tests: no

Site selection: sectioned teeth, through locations used for index test

Target condition: sound, superficial enamel demineralisation, enamel breakdown, dentine caries

Flow and timing

Participants with index test but no reference standard: 0

Participants with reference standard but no index test: 0

Time interval between tests: minimal

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Nakajima 2014 (Continued)

Was a case-control design avoided?	Yes	
Did the study avoid inappropriate exclusions?	Unclear	
Could the selection of patients have introduced bias?		High risk
Are there concerns that the included patients and setting do not match the review question?		High
DOMAIN 2: Index Test (All tests)		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
If a threshold was used, was it pre-specified?	Yes	
If multiple tests were applied were different examiners used for each (in vivo)?	No	
Could the conduct or interpretation of the index test have introduced bias?		Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

NCT02657538

Study characteristics

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: sound or early enamel lesions; "... at least one non-restored interdental surface with either an approximal enamel caries and/or a sound tooth surface in the posterior teeth"</p> <p>Teeth: permanent premolars and molars</p> <p>Sealants: not reported</p> <p>Restorations: excluded</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: minimum age of 15 years; mean age: 27.5 years, range of age: 18 to 37 years</p> <p>Sex: 11 male, 24 female</p> <p>Ethnicity: not reported</p> <p>Country: Switzerland</p> <p>Setting: dental hospital - restorative department</p> <p>Number of participants/teeth/sites: 35 participants, 70 teeth</p> <p>Prevalence: enamel 0.56, dentine 0.03</p>
Index tests	<p>Category of test: NIR - DIAGNOcam</p> <p>Sequence of test(s): radiograph and NIR performed prior to tooth separation but assessed after tooth separation, with blinding to results of reference standard</p> <p>Examiner training and calibration: 2 investigators - calibration performed for IC-DAS and DIAGNOcam</p> <p>Threshold applied: "0: Sound surfaces. 1: First visible signs of enamel caries. 2: Established caries lesion without any contact to the enamel-dentine junction (EDJ). 3: Established enamel caries with an isolated spot reaching the EDJ. 4: Dentine caries penetrating the EDJ linearly (broad contact to the EDJ). 5: Deep dentine caries lesion with visible shadow in the dentine"; "The rating of the bitewing radiographs was done from D0 to D4"</p> <p>Device specifics: air drying of teeth, no other illumination in the dental unit</p>
Target condition and reference standard(s)	<p>Category: visual - with additional impression of separated surfaces</p> <p>Sequence of index test and reference standard: NIR not likely to affect results of reference standard</p> <p>Training of examiner: same examiners as index test</p> <p>Blinding to index test: unclear</p> <p>Multiple tests: visual and radiograph performed before NIR, same examiners for all, but visual assessed before radiograph</p> <p>Site selection: all surfaces</p> <p>Target condition: ICDAS II criteria</p>
Flow and timing	<p>Participants with index test but no reference standard: 0</p>

NCT02657538 (Continued)

Participants with reference standard but no index test: 0

Time interval between tests: minimal

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Combined test data also available but not used here

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

NCT02657538 (Continued)

Could the reference standard, its conduct, or its interpretation have introduced bias?

High risk

Are there concerns that the target condition as defined by the reference standard does not match the question?

Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Could the patient flow have introduced bias?

Low risk

Obry-Musset 1988
Study characteristics

Patient Sampling	Method of sampling: unclear Included conditions: unclear Teeth: primary and permanent, premolars and molars Sealants: not reported Restorations: not reported Surface: approximal
Patient characteristics and setting	Age: 5 to 35 years Sex: not reported Ethnicity: not reported Country: France Setting: dental hospital Number of participants/teeth/sites: 330 teeth, 3960 sites (12 sites per participant) Prevalence: enamel 0.16, dentine 0.07
Index tests	Category of test: FOTI Sequence of test(s): index test prior to reference standard Examiner training and calibration: not reported Threshold applied: sound (no shadow), white or brown clinical spot (no shadow), shadow in outer dentine, shadow in inner dentine; white or brown category classed as enamel caries in our results

Obry-Musset 1988 (Continued)

	Device specifics: Oralum lamp (Rocky Mtn, Denver), 1.2 diameter
Target condition and reference standard(s)	Category: conventional radiograph Sequence of index test and reference standard: following index test Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: full surface radiographed Target condition: sound, "detectable changes, small cavity < 2 mm, large cavity > 2 mm"
Flow and timing	Participants with index test but no reference standard: 9 radiographs could not be interpreted, 14 showed orthodontic treatment Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Only permanent surfaces used in results

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Obry-Musset 1988 (Continued)

Are there concerns that the index test, its conduct, or interpretation differ from the review question?

Low concern

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?

No

Were the reference standard results interpreted without knowledge of the results of the index tests?

Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias?

High risk

Are there concerns that the target condition as defined by the reference standard does not match the question?

High

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

No

Were all patients included in the analysis?

No

Could the patient flow have introduced bias?

High risk

Park 2018
Study characteristics

Patient Sampling

Method of sampling: selected

Included conditions: unclear

Teeth: permanent incisors (25), premolars (11) and molars (15)

Sealants: not reported

Restorations: not reported

Surface: smooth

Patient characteristics and setting

Age: not reported

Sex: not reported

Ethnicity: not reported

Country: Germany

Setting: extracted teeth

Number of participants/teeth/sites: 51 teeth, 133 sites

Prevalence: enamel 0.73, dentine 0.05

Index tests

Category of test: spectral domain OCT

Park 2018 (Continued)

	Sequence of test(s): ordering of tests not clear
	Examiner training and calibration: examiner was an experienced user of OCT
	Threshold applied: ICDAS 0-4 was applied to OCT
	Device specifics: Telesto II (Thorlabs GmbH, Dachau, Germany); "Technical specifications of the SD-OCT system were as follows: center wavelength 1310 nm $\pm$ 107 nm, sensitivity $\leq$ 106 dB, axial/lateral resolution < 7.5 (air)/15 μ m, field of view $\leq$ 10 mm $\times$ 10 mm $\times$ 3.5 mm (pixel size 700 $\times$ 700 $\times$ 512), imaging speed 48–91 kHz, and A-scan average 1–5"
Target condition and reference standard(s)	Category: visual
	Sequence of index test and reference standard: unclear ordering of tests
	Training of examiner: not reported
	Blinding to index test: unclear
	Multiple tests: no
	Site selection: specified regions of interest
	Target condition: ICDAS 0-4
Flow and timing	Participants with index test but no reference standard: 0
	Participants with reference standard but no index test: 0
	Time interval between tests: minimal
	Participants receiving both tests but excluded from results: 0
Comparative	
Notes	3 studies in this paper that could have been included. The first was in vitro, OCT (swept source) versus histology, could not be included as no data were presented. The second, in vitro part (OCT - spectral domain) is included, above. The third in vivo element (OCT - spectral domain) but samples participants with caries up to ICDAS 4, so we had to exclude this study

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	

Park 2018 (Continued)

Are there concerns that the included patients and setting do not match the review question? High

DOMAIN 2: Index Test (All tests)

Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

If a threshold was used, was it pre-specified? Yes

If multiple tests were applied were different examiners used for each (in vivo)?

Could the conduct or interpretation of the index test have introduced bias? Unclear risk

Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? No

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Could the patient flow have introduced bias? Low risk

Shimada 2010
Study characteristics

Patient Sampling	Method of sampling: selected
	Included conditions: "stained occlusal fissures and/or small open caries lesions with diameter within 1 mm"; "36 non-cavitated teeth and 26 cavitated teeth" - not clear on severity of cavitation, concern that no obviously sound surfaces were sampled

Shimada 2010 (Continued)

	Teeth: permanent molars Sealants: not reported Restorations: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Japan Setting: extracted teeth Number of participants/teeth/sites: 62 teeth, 111 sites Prevalence: enamel 0.86, dentine 0.38
Index tests	Category of test: SS-OCT Sequence of test(s): visual test before OCT (examiners blinded to results of visual test), both prior to reference standard Examiner training and calibration: 3 experienced dentists, with training session for calibration Threshold applied: sound, superficial demineralisation of enamel, enamel breakdown due to the caries, dentine caries Device specifics: SS-OCT (Santec OCT-20001, Santec Co, Komaki, Japan); "Axial resolution of the system is 11 mm in air, which corresponds to 8 mm in soft tissue and 6.8 mm in enamel, assuming refractive indices of about 1.38 and 1.62, respectively"
Target condition and reference standard(s)	Category: histology - using confocal laser scanning microscope Sequence of index test and reference standard: index test then reference standard Training of examiner: "Two examiners with sufficient experience in histopathological study of caries" Blinding to index test: unclear Multiple tests: no Site selection: "performed on the investigation site" Target condition: sound, superficial enamel demineralisation, enamel breakdown, dentine caries
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	

Shimada 2010 (Continued)

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Shimada 2010 (Continued)

Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Shimada 2014

Study characteristics

Patient Sampling	<p>Method of sampling: selected; "Ninety-one unrestored proximal surfaces of premolars and molars, with/without the possibility of caries, were selected for this study"</p> <p>Included conditions: sound and early lesions</p> <p>Teeth: permanent molar and premolars</p> <p>Sealants: not reported</p> <p>Restorations: no</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: 21 to 64 years (mean 38 years)</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Japan</p> <p>Setting: Tokyo Medical and Dental University</p> <p>Number of participants/teeth/sites: 53 participants, 91 surfaces</p> <p>Prevalence: enamel 0.67, dentine 0.38</p>
Index tests	<p>Category of test: SS-OCT</p> <p>Sequence of test(s): visual test before OCT (examiners blinded to results of visual test), both prior to reference standard</p> <p>Examiner training and calibration: 6 dentists, with training session for calibration</p> <p>Threshold applied: sound, superficial demineralisation of enamel, enamel breakdown due to the caries, dentine caries</p> <p>Device specifics: "Dental OCT System (Prototype 2, Panasonic Healthcare Co Ltd, Ehime, Japan) used in the present study"; "high-speed frequency swept laser light with a center wavelength of 1330 nm was projected onto the occlusal surface"</p>
Target condition and reference standard(s)	<p>Category: clinical visual evaluation - excavation of severe caries, the remainder were based on a visual examination (no separation of teeth but only those that could be viewed were included)</p>

Shimada 2014 (Continued)

Sequence of index test and reference standard: index test prior to reference standard

Training of examiner: not reported

Blinding to index test: not possible

Multiple tests: yes, visual plus excavation

Site selection: approximal surface under investigation

Target condition: no caries, enamel, or dentine

Flow and timing

Participants with index test but no reference standard: 5 excluded from results since the approximal surface could not be viewed

Participants with reference standard but no index test: 0

Time interval between tests: unclear

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Shimada 2014 (Continued)

Are there concerns that the index test, its conduct, or interpretation differ from the review question?

Low concern

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? No

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Unclear

Did all patients receive the same reference standard? No

Were all patients included in the analysis? No

Could the patient flow have introduced bias? High risk

Sidi 1988

Study characteristics

Patient Sampling	<p>Method of sampling: selected - participants taking part in a caries trial</p> <p>Included conditions: only surfaces considered to be sound by clinical evaluation were evaluated with the transillumination device</p> <p>Teeth: permanent molar and premolars</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: 21 to 64 years (mean 28 years)</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: not reported</p> <p>Setting: clinical trial; "carried out away from the dental surgery"</p>

Sidi 1988 (Continued)

	<p>Number of participants/teeth/sites: 456 participants, 4405 surfaces (71 participants could not be radiographed within the study)</p> <p>Prevalence: enamel 0.03, dentine 0.01 (very low prevalence results in high concern for applicability)</p>
Index tests	<p>Category of test: FOTI</p> <p>Sequence of test(s): visual examination preceded radiographs and transillumination, radiographs developed and interpreted at a later date</p> <p>Examiner training and calibration: not reported</p> <p>Threshold applied: sound, enamel, dentine</p> <p>Device specifics: liquid light transmission cord: Kulzer and Co, placed buccally and lingually</p>
Target condition and reference standard(s)	<p>Category: radiographs</p> <p>Sequence of index test and reference standard: reference standard completed prior to transillumination</p> <p>Training of examiner: not reported</p> <p>Blinding to index test: yes - attempt made to blind examiners to index test results</p> <p>Multiple tests: visual completed but without separation so not a viable reference standard</p> <p>Site selection: all surfaces</p> <p>Target condition: sound, enamel, or dentine</p>
Flow and timing	<p>Participants with index test but no reference standard: all participants that received radiographs also were examined with transillumination - 71 participants did not receive radiographs, unclear on the number that received a clinical examination but did not receive transillumination test</p> <p>Participants with reference standard but no index test: 0</p> <p>Time interval between tests: unclear</p> <p>Participants receiving both tests but excluded from results: 0</p>
Comparative	
Notes	Buccal and lingual results totaled for our data analysis, from table IV

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		

Sidi 1988 (Continued)

Could the selection of patients have introduced bias?	High risk	
Are there concerns that the included patients and setting do not match the review question?		High
DOMAIN 2: Index Test (All tests)		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
If a threshold was used, was it pre-specified?	Yes	
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear	
Could the conduct or interpretation of the index test have introduced bias?	Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	No	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?		High
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Unclear	
Were all patients included in the analysis?	Unclear	
Could the patient flow have introduced bias?	Unclear risk	

Simon 2016
Study characteristics

Simon 2016 (Continued)

Patient Sampling	<p>Method of sampling: systematically sampled from attending population, no clear reporting of inclusion criteria regarding the caries status in sampled teeth</p> <p>Included conditions: participants scheduled for orthodontic extractions</p> <p>Teeth: permanent premolars</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: occlusal and approximal</p>
Patient characteristics and setting	<p>Age: 12 to 60 years</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: USA</p> <p>Setting: clinical setting, dental school, followed by extraction</p> <p>Number of participants/teeth/sites: 40 participants teeth, 109 occlusal and approximal sites</p> <p>Prevalence:</p> <p>occlusal - enamel 0.82, dentine not reported</p> <p>approximal - enamel 0.28, dentine not reported</p>
Index tests	<p>Category of test: NIR transillumination</p> <p>Sequence of test(s): index tests in the clinical setting in the following order: (i) conventional photos; (ii) cross-polarized NIR; (iii) occlusal NIR; (iv) approximal NIR. Digital radiographs followed in the in vitro setting</p> <p>Examiner training and calibration: 1 examiner assessed the NIR and another the radiographs, each was an experienced dentist with over 20 years experience</p> <p>Threshold applied: S - sound; E1 - outer half of enamel; E2 - inner half of enamel; D1 - inner half of dentine; D2 - second half of dentine</p> <p>Device specifics: "Light centered at 1310 nm is generated using a superluminescent laser diode (SLD), Model SLD72 (COVEGA Corporation, Jessup, MD) with 50 nm bandwidth. Fiber optic cables are used to deliver the light into a Teflon diffusing element"</p> <p>Note - high applicability as device is not commercially available; "near-IR transillumination and near-IR reflectance probes that we fabricated in our laboratory for both occlusal and approximal lesions"</p>
Target condition and reference standard(s)	<p>Category: histology - polarized light microscopy (PLM) was used for histological examination</p> <p>Sequence of index test and reference standard: index test then reference standard</p> <p>Training of examiner: not reported</p> <p>Blinding to index test: unclear</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p>

Simon 2016 (Continued)

Target condition: sound, inner/outer enamel, inner/outer dentine

Flow and timing

Participants with index test but no reference standard: 0

Participants with reference standard but no index test: 0

Time interval between tests: minimal

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Simon 2016 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Simon 2016a

Study characteristics	
Patient Sampling	Added to generate data entry for approximal surfaces
Patient characteristics and setting	
Index tests	
Target condition and reference standard(s)	
Flow and timing	
Comparative	
Notes	

Van Hilsen 2013

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: sound and non-cavitated; "A single examiner sorted through collected teeth and chose an assortment of teeth without evidence of cavitated lesions (ICDAS-II 0–2)" Teeth: permanent posterior teeth Sealants: excluded

Van Hilsen 2013 (Continued)

	Restorations: excluded
	Surface: occlusal
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: USA</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 45 teeth/sites</p> <p>Prevalence: enamel 0.76, dentine 0.31</p>
Index tests	<p>Category of test: CP-OCT; "CP-OCT images were presented to two blinded examiners who had at least 4 months experience assessing CP-OCT images of sound, non-cavitated, and cavitated lesions. Examiners (E5, E6) independently graded the images"</p> <p>Sequence of test(s): MidWest caries, visual, photographic, OCT prior to reference standard - all blinded with different assessors</p> <p>Examiner training and calibration: 2 blinded examiners with at least 4 months experience of OCT</p> <p>Threshold applied: sound, incipient subsurface (enamel), extensive subsurface (dentine)</p> <p>Device specifics: "cross-polarization swept source OCT (CP-OCT) system with an intraoral probe (IVS-200-CPM, Santec Co Komaki, Japan)"</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: index test then reference standard</p> <p>Training of examiner: not reported</p> <p>Blinding to index test: yes</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: sound, enamel, dentine</p>
Flow and timing	<p>Participants with index test but no reference standard: 3 damaged during sectioning of tooth</p> <p>Participants with reference standard but no index test: 0</p> <p>Time interval between tests: minimal</p> <p>Participants receiving both tests but excluded from results: 0</p>
Comparative	
Notes	

Van Hilsen 2013 (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Van Hilsen 2013 (Continued)

Were all patients included in the analysis? No

Could the patient flow have introduced bias? Low risk

Xiao-Hua 2016
Study characteristics

Patient Sampling	Method of sampling: selected Included conditions: sound and non-cavitated Teeth: permanent Sealants: not reported Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: China Setting: extracted teeth Number of participants/teeth/sites: 77 teeth (97 reported in results) Prevalence: enamel 0.74, dentine 0.48
Index tests	Category of test: high-speed OCT Sequence of test(s): visual and OCT, all prior to reference standard Examiner training and calibration: 3 examiners reached a consensus, training unclear Threshold: sound, enamel, dentine Device specifics: "wavelength 1310 nm, bandwidth 50 nm, coherence length 15 µm, resolution within the tooth sample is 10 µm"
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth

Xiao-Hua 2016 (Continued)

Flow and timing

Target condition: sound, enamel, dentine

Participants with index test but no reference standard: 0

Participants with reference standard but no index test: 0

Time interval between tests: minimal

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Xiao-Hua 2016 (Continued)

Are there concerns that the target condition as defined by the reference standard does not match the question?

Low concern

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Could the patient flow have introduced bias?

Low risk

Zain 2018
Study characteristics

Patient Sampling	<p>Method of sampling: selected; "ninety preliminary investigation sites with IC-DAS code 0, 1, 2 (n = 30 each) were identified"</p> <p>Included conditions: sound and early caries teeth</p> <p>Teeth: permanent premolars</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: occlusal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Malaysia</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 54 teeth, 71 sites</p> <p>Prevalence: enamel 0.7, dentine 0</p>
Index tests	<p>Category of test: SS-OCT Imaging System (OCS1300SS, Thorlabs Ltd, UK)</p> <p>Sequence of test(s): OCT prior to reference standard</p> <p>Examiner training and calibration: "two trained and calibrated examiners"</p> <p>Threshold applied: "non-cavitated fissure caries is considered to be present when bands of elevated back scattered intensity, in the range of -15 dB to -5 dB, which are sub-surface and diffused in nature were observed. Intensity in the same range which is limited to one to two pixels thick on the surface was NOT considered as carious"</p> <p>Device specifics: "intensity range is presented as yellow and red by the Large DR3 colour map of the Thorlabs OCT capturing software"</p>

Zain 2018 (Continued)

Target condition and reference standard(s)	<p>Category: histology - polarized light microscopy</p> <p>Sequence of index test and reference standard: index test then reference standard</p> <p>Training of examiner: 2 trained and calibrated examiners</p> <p>Blinding to index test: not clearly reported</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: sound, enamel caries; "absence of dark brown and black areas at the investigation sites were considered as sound whereas, the presence of these dark brown and black areas were regarded as carious"</p>
Flow and timing	<p>Participants with index test but no reference standard: unclear "a final cohort of seventy one investigation sites were selected based on the Ekstrand histology criteria"</p> <p>Participants with reference standard but no index test: 0</p> <p>Time interval between tests: minimal</p> <p>Participants receiving both tests but excluded from results: 71/104 included in final analysis</p>

Comparative

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			

Zain 2018 (Continued)

Could the conduct or interpretation of the index test have introduced bias?	Low risk
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	High risk

CP-OCT = cross-polarization optical coherence tomography; DIFOTI = digital fibre-optic transillumination; ECM = electronic caries monitor; FOTI = fibre-optic transillumination; ICDAS = International Caries Detection and Assessment System; microCT = microcomputed tomography; NIR = near-infrared; OCT = optical coherence tomography; QLF = quantitative light-induced fluorescence; SS-OCT = swept source optical coherence tomography.

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Baltacioglu 2017	Data not available to produce a 2 x 2 table at the enamel caries threshold
Bin-Shuwaish 2008	Participant sample included patients requiring restoration
Blazejewska 2016	Reference standard was visual with ICDAS but without tooth separation - this is not acceptable as any lesions cannot be directly viewed, so excluded
Chesters 2002	Not possible to complete a 2 x 2 table of results at the enamel threshold
Choksi 1994	No reference standard
Cortes 2000	Sample included dentinal lesions

Study	Reason for exclusion
Cortes 2003	Sample included dentinal lesions
da Silva 2011	Not possible to complete a 2 x 2 table of results
Elhennawy 2018	Not possible to complete a 2 x 2 table of results at the enamel threshold
Gomez 2013	Included teeth with ICDAS 4, therefore dentine caries
Grossman 2002	Sample included dentinal lesions
Hintze 1998	Not possible to complete a 2 x 2 table of results at the enamel threshold
Holtzman 2014	Sample included ICDAS 4 and we could not extract a 2 x 2 table of results at the enamel threshold
Holtzman 2015	Sample included ICDAS 4 and we could not extract a 2 x 2 table of results at the enamel threshold
Ibusuki 2015	Not possible to complete a 2 x 2 table of results at the enamel threshold
Kuhnisch 2016	Sample included dentinal lesions
Lara-Capi 2017	No suitable reference standard
Lederer 2019	Inappropriate sample - selected teeth with dentinal caries (ICDAS 4 and 5)
Lee 2009	Not possible to complete a 2 x 2 table of results at the enamel threshold
Lee 2010	Not possible to complete a 2 x 2 table of results at the enamel threshold
Lee 2010a	Not possible to complete a 2 x 2 table of results at the enamel threshold
Manton 2007	Not possible to complete a 2 x 2 table of results at the enamel threshold
Marianska 2015	Inappropriate sample - selected teeth with dentinal caries which were due restorations
Marinova-Takorova 2014	Not a DTA study. No sound teeth included
Matsuura 2018	Not possible to complete a 2 x 2 table of results at the enamel threshold
Mialhe 2009	Same study as Mialhe 2003 but sound surfaces removed from the results
Mitropoulos 1985	Diagnostic threshold applied at the dentine level only
Mitropoulos 1985a	Diagnostic threshold applied at the dentine level only, so a 2 x 2 table at enamel threshold impossible to obtain
Ozkan 2017	No sound teeth included in sample
Purdell-Lewis 1974	Not possible to complete a 2 x 2 table of results at the enamel threshold
Salsone 2012	Sample included dentinal lesions
Sasazawa 2015	Not possible to complete a 2 x 2 table of results at the enamel threshold
Schaefer 2018	Not possible to complete a 2 x 2 table of results at the enamel threshold due to reporting in study results of sealant, restorations and not assessable categories

Study	Reason for exclusion
Schneiderman 1997	Sample included teeth with frank cavitation
Simon 2014	Not possible to complete a 2 x 2 table of results at the enamel threshold
Simon 2018	Not possible to complete a 2 x 2 table of results at the enamel threshold, reported lesion depth in mm
Sochtig 2014	Sample included dentinal lesions
Staninec 2010	Not possible to complete a 2 x 2 table of results at the enamel threshold
Staninec 2011	Sample included dentinal lesions
Stephen 1987	Appropriate sample, index test, and reference standard but data not available at the enamel threshold, table 3 presented sound versus codes 3/4 which is caries in dentine or pulp or both
Tetschke 2018	Not possible to complete a 2 x 2 table of results at the enamel threshold
Verdonschot 1991	Not possible to complete a 2 x 2 table since prevalence from histology was not reported
Verdonschot 1992	Sample included dentinal lesions
Waly 1995	Not possible to complete a 2 x 2 table of results at the enamel threshold
Wenzel 1992	Sample included dentinal lesions
Wilder-Smith 2013	Abstract only, cannot create a 2 x 2 table at the enamel threshold
Yu 2017	Not possible to complete a 2 x 2 table of results at the enamel threshold
Zakian 2009	Not possible to complete a 2 x 2 table of results at the enamel threshold

DTA = diagnostic test accuracy; ICDAS = International Caries Detection and Assessment System.

DATA

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 ALL	24	16702
2 NIR	6	673
3 OCT	10	1171
4 FOTI/DIFOTI	8	14858
5 DIFOTI	2	2200

Test 1. ALL

ALL

Test 2. NIR

NIR

Test 3. OCT

OCT

Test 4. FOTI/DIFOTI

FOTI/DIFOTI

Study	TP	FP	FN	TN	Sensitivity {95% CI}	Specificity {95% CI}	Sensitivity {95% CI}	Specificity {95% CI}
Ashley 1998	13	5	49	36	0.21 [0.12, 0.33]	0.88 [0.74, 0.96]
Astvaldsdottir 2012	41	3	22	31	0.65 [0.52, 0.77]	0.91 [0.76, 0.98]
Chawla 2012	30	2	52	51	0.37 [0.26, 0.48]	0.96 [0.87, 1.00]
Holt 1989	41	33	20	687	0.67 [0.54, 0.79]	0.95 [0.94, 0.97]
Laitala 2017	225	410	190	1278	0.54 [0.49, 0.59]	0.76 [0.74, 0.78]
Mialhe 2003	65	8	106	20	0.38 [0.31, 0.46]	0.71 [0.51, 0.87]
Obry-Musset 1988	282	99	146	2118	0.66 [0.61, 0.70]	0.96 [0.95, 0.96]
Sidi 1988	73	179	165	8378	0.31 [0.25, 0.37]	0.98 [0.98, 0.98]

Test 5. DIFOTI

DIFOTI

Study	TP	FP	FN	TN	Sensitivity {95% CI}	Specificity {95% CI}	Sensitivity {95% CI}	Specificity {95% CI}
Astvaldsdottir 2012	41	3	22	31	0.65 [0.52, 0.77]	0.91 [0.76, 0.98]
Laitala 2017	225	410	190	1278	0.54 [0.49, 0.59]	0.76 [0.74, 0.78]

ADDITIONAL TABLES

Table 1. Classification of levels of caries levels

DMFT classification	Definition (Pitts 2001)
0	Sound (non-diseased)
D ₁	Non-cavitated yet clinically detectable enamel lesions with intact surfaces
D ₂	Cavitated lesion penetrating the enamel or shadowing
D ₃	Cavity progressing past the enamel-dentine junction into dentine
D ₄	Cavity progressing into pulp

DMFT = decayed, missing, and filled teeth.

Table 2. Index tests for caries

Test	Characteristics	Intended use in clinical pathway	Other information
Visual or visual-tactile examination	Identifying caries according to their visual appearance, aided by a dental mirror and probe, on clean and dry teeth	The fundamental step in the detection of caries, but limited in the diagnosis of early lesions. All patients presenting to a dental clinician will receive a visual examination	<p>Advantages: completed and interpreted quickly with minimal invasion and little cost except clinician training and time</p> <p>Disadvantages: early caries are difficult to observe visually, depth and severity of lesions cannot be as-</p>

Table 2. Index tests for caries (Continued)

			<p>essed, approximal lesions cannot be seen</p>
Radiography	<p>Bitewing radiology is the most commonly used method. Others include: subtraction radiographs which provides a semi-automated method for monitoring progression of lesions (Ellwood 1997; Wenzel 2000) and cone beam computed technology (CBCT) which provides a 3-dimensional image which appears to offer great potential for diagnosis with increased levels of radiation (Horner 2009)</p>	<p>Widely used as an adjunct to aid detection and in particular to inform the clinician of the depth and severity of lesion (Wenzel 1995; Whaites 2013)</p> <p>Relevant on occlusal surfaces but also in approximal location which are otherwise difficult to assess visually</p>	<p>Advantages: radiographs aid the detection of caries and are shown to be more sensitive than visual examination on approximal and occlusal lesions (Wenzel 2004)</p> <p>Disadvantages: limitations exist when detecting early caries in enamel surfaces. There is a small but real risk over patient exposure to ionising radiation, which has to be balanced with the patient's age, caries risk, and time since previous radiograph (Pitts 2017). Digital radiographic methods have shown benefits for patients with the speed in which they can be viewed and for the ability to manipulate images for increased clarity (Wenzel 2006)</p>
Fluorescence	<p>The breakdown of enamel alters the characteristics of its structure, when exposed to light-inducing fluorescence diseased teeth respond differently to sound teeth. There is potential for mineral loss to be quantified and used to aid the diagnostic decision and treatment pathway (Angmar-Månsson 2001; Matos 2011). Fluorescence is typically divided into laser fluorescence and light fluorescence (i.e. DIAGNOdent type devices and quantitative light-induced fluorescence (QLF) type devices)</p>	<p>Potential to aid the clinician in identifying early caries which may not be possible with a visual examination alone. QLF emits either green or red light and may ascertain whether the lesion is active or arrested</p>	<p>Advantages: the potential to identify changes in tooth characteristics that are otherwise unobservable in a visual-tactile examination</p> <p>Disadvantages: uncertainty of the reliability of devices and the ability to detect disease and health</p>
Fibre-optic transillumination	<p>Fibre-optic transillumination (FOTI) uses a light emitted from a handheld device which when placed directly onto the tooth illuminates the tooth (Pretty 2006). Any demineralisation should appear as shadows in the tooth due to the disruption of the tooth's structure due to caries</p>	<p>An adjunct to the visual examination, particularly useful for detecting approximal caries, with its strength being in identifying early caries in enamel and dentine (Davies 2001). A further advancement with fibre-optic techniques combines this with a camera to capture an image which may or may not be linked to software for analysis, digital imaging FOTI (DIFOTI)</p>	<p>Advantages: the potential to identify changes in tooth characteristics that are otherwise unobservable in a visual-tactile examination</p> <p>Disadvantages: uncertainty of the reliability of devices and the ability to detect disease and health</p>
Electrical conductance	<p>The demineralisation of the tooth is reported to have an effect on the tooth's electrical conductance. This is measured by placing a probe on the tooth which measures any potentially higher conductivity which occurs due to carious lesions being filled with saliva (Tam 2001)</p>	<p>An adjunct to the visual examination</p>	<p>Advantages: the potential to identify changes in tooth characteristics that are otherwise unobservable in a visual-tactile examination</p> <p>Disadvantages: uncertainty of the reliability of devices and the ability to detect disease and health. Particularly due to the necessity to place the</p>

Table 2. Index tests for caries (Continued)

probe in an identical location for a reproducible result

Table 3. QUADAS-2 tool

Item	Response (delete as required)
Participant selection – Risk of bias	
1) Was a consecutive or random sample of participants or teeth used?	<p>Yes – where teeth or participants were selected consecutively or allocated to the study via a randomisation process</p> <p>No – if study described another method of sampling</p> <p>Unclear – if participant sampling is not described</p>
2) Was a case-control design avoided?	<p>Yes – if case-control clearly not used</p> <p>No – if study described as case-control or describes sampling specific numbers of participants with particular diagnoses</p> <p>Unclear – if not clearly described</p>
3) Did the study avoid inappropriate exclusions (e.g. inclusion of caries into dentine)?	<p>Yes – if the study clearly reports that included participants or teeth were apparently healthy or caries into dentine were excluded</p> <p>No – if lesions were included that showed caries into dentine or exclusions that might affect test accuracy (e.g. teeth with no caries)</p> <p>Unclear – if not clearly reported</p>
Could the selection of participants have introduced bias?	
If answers to all of questions 1) and 2) and 3) was 'yes'	Risk is Low
If answers to any of questions 1) and 2) and 3) was 'no'	Risk is High
If answers to any of questions 1) and 2) and 3) was 'unclear'	Risk is Unclear
Participant selection – Concerns regarding applicability	
1) Does the study report results for participants or teeth selected by apparent health or suspected early caries (i.e. studies do not recruit patients who are known to have advanced caries into dentine)?	<p>Yes – if a group of participants or teeth has been included which is apparently healthy or indicative of early caries</p> <p>No – if a group of participants or teeth has been included which is suspected of advanced caries</p> <p>Unclear – if insufficient details are provided to determine the spectrum of participants or teeth</p>
2) Did the study report data on a per-patient rather than on a tooth or surface basis?	<p>Yes – if the analysis was reported on a surface or tooth basis</p> <p>No – if the analysis was reported on a per-patient basis</p>

Table 3. QUADAS-2 tool (Continued)

	Unclear - if it is not possible to assess whether data are presented on a per-patient or per-tooth basis
3) Did the study avoid an in vitro setting which required the usage of extracted teeth?	Yes – if the participants were recruited prior to tooth extraction No – if previously extracted teeth were used in the analysis Unclear – if it was not possible to assess the source and method of recruiting of included participants/teeth
Is there concern that the included participants or teeth do not match the review question?	
If answers to all of questions 1) and 2) and 3) was 'yes'	Risk is Low
If answers to any of questions 1) and 2) and 3) was 'no'	Risk is High
If answers to any of questions 1) and 2) and 3) was 'unclear'	Risk is Unclear
Index test - Risk of bias (to be completed per test evaluated)	
1) Was the index test result interpreted without knowledge of the results of the reference standard?	Yes – if the index test described is always conducted and interpreted prior to the reference standard result, or for retrospective studies interpreted without prior knowledge of the reference standard No – if index test described as interpreted in knowledge of reference standard result Unclear – if index test blinding is not described
2) Was the diagnostic threshold at which the test was considered positive pre-specified?	Yes – if threshold was pre-specified (i.e. prior to analysing the study results) No – if threshold was not pre-specified Unclear – if not possible to tell whether or not diagnostic threshold was pre-specified
For visual and radiograph tests only: 3) For studies reporting the accuracy of multiple diagnostic thresholds for the same index test or multiple index tests, was each threshold or index test interpreted without knowledge of the results of the others?	Yes – if thresholds or index tests were selected prospectively and each was interpreted by a different clinician or interpreter, or if study implements a retrospective (or no) cut-off (i.e. look for deepest/most severe lesion first) No – if study states reported by same reader Unclear - if no mention of number of readers for each threshold or if pre-specification of threshold not reported N/A - multiple diagnostic thresholds not reported for the same index test
Could the conduct or interpretation of the index test have introduced bias?	
For visual and radiographic studies item 3) to be added	
If answers to all of questions 1) and 2) was 'yes'	Risk is Low
If answers to any of questions 1) and 2) was 'no'	Risk is High

Table 3. QUADAS-2 tool (Continued)

If answers to any of questions 1) and 2) was 'unclear'	Risk is Unclear
Index test - Concerns regarding applicability	
1) Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	<p>Yes – if the criteria for detection or diagnosis of the target disorder were reported in sufficient detail to allow replication</p> <p>No – if the criteria for detection or diagnosis of the target disorder were not reported in sufficient detail to allow replication</p> <p>Unclear – if some but not sufficient information on criteria for diagnosis to allow replication were provided</p>
2) Was the test interpretation carried out by an experienced examiner?	<p>Yes – if the test clearly reported that the test was interpreted by an experienced examiner</p> <p>No – if the test was not interpreted by an experienced examiner</p> <p>Unclear – if the experience of the examiner(s) was not reported in sufficient detail to judge or if examiners described as 'Expert' with no further detail given</p>
Is there concern that the included participants do not match the review question?	
If the answer to question 1) and 2) was 'yes'	Concern is Low
If the answer to question 1) and 2) was 'no'	Concern is High
If the answer to question 1) and 2) was 'unclear'	Concern is Unclear
Reference standard - Risk of bias	
1) Is the reference standard likely to correctly classify the target condition?	<p>Yes – if all teeth or surfaces underwent a histological or excavation reference standard</p> <p>No – if a final diagnosis for any participant or tooth was reached without the histological or excavation reference standards</p> <p>Unclear – if the method of final diagnosis was not reported</p>
2) Were the reference standard results interpreted without knowledge of the results of the index test?	<p>Yes – if the reference standard examiner was described as blinded to the index test result</p> <p>No – if the reference standard examiner was described as having knowledge of the index test result</p> <p>Unclear – if blinded reference standard interpretation was not clearly reported</p>
Could the reference standard, its conduct, or its interpretation have introduced bias?	
If answers to questions 1) and 2) was 'yes'	Risk is Low
If the answer to question 1) and 2) was 'no'	Concern is High
If the answer to question 1) and 2) was 'unclear'	Concern is Unclear
Reference standard - Concerns regarding applicability	

Table 3. QUADAS-2 tool (Continued)

1) Does the study use the same definition of disease positive as the prescribed in the review question?	<p>Yes - same definition of disease positive used, or teeth can be disaggregated and re-grouped according to review definition</p> <p>No - some teeth cannot be disaggregated</p> <p>Unclear - definition of disease positive not clearly reported</p>
Flow and timing - Risk of bias	
1) Was there an appropriate interval between index test and reference standard (in vivo studies less than 3 months, in vitro no limit but must be stored appropriately)?	<p>Yes - if study reports index and reference standard had a suitable interval or storage method</p> <p>No - if study reports greater than 3-month interval between index and reference standard or inappropriate storage of extracted teeth prior to reference standard</p> <p>Unclear - if study does not report interval or storage methods between index and histological reference standard</p>
2) Did all participants receive the same reference standard?	<p>Yes - if all participants underwent the same reference standard</p> <p>No - if more than 1 reference standard was used</p> <p>Unclear - if not clearly reported</p>
3) Were all participants included in the analysis?	<p>Yes - if all participants were included in the analysis</p> <p>No - if some participants were excluded from the analysis</p> <p>Unclear - if not clearly reported</p>
If answers to questions 1) and 2) and 3) was 'yes'	Risk is Low
If answers to any one of questions 1) or 2) or 3) was 'no'	Risk is High
If answers to any one of questions 1) or 2) or 3) was 'unclear'	Risk is Unclear

N/A = not applicable; QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies 2.

APPENDICES

Appendix 1. MEDLINE Ovid search strategy

- exp Tooth demineralization/
- (teeth adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
- (tooth adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
- (dental adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
- (enamel adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
- (dentin adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
- (root adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
- or/1-7
- fiber optic technology/
- optical fibers/
- ((fiber or fibre) adj optic\$).mp.
- (FOTI or DiFOTI).mp.
- transillumina\$.mp.
- (DiagnoCAM or CariVu).mp.

15. ("near infra red" or "near infrared" or NIR or N-I-R or "near ir").mp.
16. Tomography, optical coherence/
17. ("optical coherence tomograph\$" or OCT or "OC tomography").ti,ab.
18. or/9-17
19. 8 and 18

Appendix 2. Embase Ovid search strategy

1. Dental caries/
2. (caries or carious).mp.
3. (teeth adj5 (cavit\$ or decay\$ or lesion\$ or deminerali\$ or reminerali\$ or fissure\$)).mp.
4. (tooth adj5 (cavit\$ or decay\$ or lesion\$ or deminerali\$ or reminerali\$ or fissure\$)).mp.
5. (dental adj5 (cavit\$ or decay\$ or lesion\$ or deminerali\$ or reminerali\$ or fissure\$)).mp.
6. (enamel adj5 (cavit\$ or decay\$ or lesion\$ or deminerali\$ or reminerali\$ or fissure\$)).mp.
7. (dentin\$ adj5 (cavit\$ or decay\$ or lesion\$ or deminerali\$ or reminerali\$ or fissure\$)).mp.
8. or/1-7
9. fiber optic technology/
10. ((fiber or fibre) adj optic\$).mp.
11. (FOTI or DiFOTI).mp.
12. transilluminate\$.mp.
13. (DiagnoCAM or CariVu).mp.
14. ("near infra red" or "near infrared" or NIR or N-I-R or "near ir").mp.
15. Tomography, optical coherence/
16. ("optical coherence tomograph\$" or OCT or "OC tomography").ti,ab.
17. or/9-16
18. 8 and 17

Appendix 3. US National Institutes of Health Ongoing Trials Register (ClinicalTrials.gov) search strategy

Expert Search interface: (caries OR tooth decay OR dental decay OR carious) AND ("fiber optic" OR "fibre optic" OR FOTI OR DiFOTI OR transillumination OR transilluminate OR diagnocam OR carivu OR "near infra red" OR "near infra-red" OR "near infrared" OR "near ir" OR "optical coherence tomography")

Appendix 4. World Health Organization International Clinical Trials Registry Platform search strategy

caries AND "fibre optic" OR caries AND "fiber optic" OR caries AND FOTI OR caries AND DiFOTI OR caries AND transilluminate OR caries AND transillumination OR caries AND diagnocam OR caries AND carivu OR caries AND "optical coherence tomography" OR caries AND "near infra red" OR caries AND "near ir" OR caries AND "near infrared"

HISTORY

Review first published: Issue 1, 2021

CONTRIBUTIONS OF AUTHORS

All review authors collaborated in the conception of the review purpose, design, and interpretation of results.

Drafting the final draft of the review: Richard Macey (RM), Tanya Walsh (TW), and Richard Hogan (RH).

Developing the search strategy: TW and RM.

Co-ordination of contributions from the co-authors: RM.

Screening of papers against eligibility criteria: RM, TW, Philip Riley (PR), Helen Worthington (HW), and Anne-Marie Glenny (AMG).

Obtained data on published, ongoing, and unpublished studies: RM and PR.

Appraising the quality of papers: RM, TW, PR, HW, and AMG.

Extracting data for the review: RM, TW, PR, HW, and AMG.

Entering data into Review Manager 5 ([Review Manager 2020](#)): RM.

Analysis of data: RM and TW.

Provided clinical guidance during all phases of review: RH, Janet Clarkson (JC), and David Ricketts (DR).

DECLARATIONS OF INTEREST

Richard Macey: none known.

Tanya Walsh: none known. I am an Editor with Cochrane Oral Health.

Philip Riley: none known. I am Deputy Co-ordinating Editor of Cochrane Oral Health.

Richard Hogan: I am employed by the Colgate-Palmolive Company as a Clinical Study Manager. I also hold an Honorary Research Fellow position with The University of Manchester.

Anne-Marie Glenney: none known. I am Co-ordinating Editor of Cochrane Oral Health.
Helen V Worthington: none known. I am Emeritus Co-ordinating Editor of Cochrane Oral Health.
Janet E Clarkson: none known. I am Co-ordinating Editor of Cochrane Oral Health.
David Ricketts: none known.

SOURCES OF SUPPORT

Internal sources

- Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, UK
- Manchester Academic Health Sciences Centre (MAHSC) and the NIHR Manchester Biomedical Research Centre, UK

External sources

- National Institute for Health Research (NIHR), UK

NIHR Cochrane Programme Grant: 16/114/23 Detection and diagnosis of common oral diseases: diagnostic test accuracy of tests of oral cancer and caries

- NIHR, UK

This project was supported by the NIHR, via Cochrane Infrastructure funding to Cochrane Oral Health. The views and opinions expressed herein are those of the review authors and do not necessarily reflect those of the Evidence Synthesis Programme, the NIHR, the NHS, or the Department of Health and Social Care

- Cochrane Oral Health Global Alliance, Other

The production of Cochrane Oral Health reviews has been supported financially by our Global Alliance since 2011 (oralhealth.cochrane.org/partnerships-alliances). Contributors in the last 2 years have been the American Association of Public Health Dentistry, USA; AS-Akademie, Germany; the British Association for the Study of Community Dentistry, UK; the British Society of Paediatric Dentistry, UK; the Canadian Dental Hygienists Association, Canada; the Centre for Dental Education and Research at All India Institute of Medical Sciences, India; the National Center for Dental Hygiene Research & Practice, USA; New York University College of Dentistry, USA; and Swiss Society of Endodontology, Switzerland