



**Cochrane**  
**Library**

Cochrane Database of Systematic Reviews

## Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)

Macey R, Walsh T, Riley P, Glenny AM, Worthington HV, O'Malley L, Clarkson JE, Ricketts D

Macey R, Walsh T, Riley P, Glenny A-M, Worthington HV, O'Malley L, Clarkson JE, Ricketts D.  
Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries.  
*Cochrane Database of Systematic Reviews* 2021, Issue 6. Art. No.: CD014546.  
DOI: [10.1002/14651858.CD014546](https://doi.org/10.1002/14651858.CD014546).

[www.cochranelibrary.com](http://www.cochranelibrary.com)

## TABLE OF CONTENTS

ABSTRACT .....	1
PLAIN LANGUAGE SUMMARY .....	2
SUMMARY OF FINDINGS .....	4
BACKGROUND .....	8
Figure 1. ....	10
OBJECTIVES .....	11
METHODS .....	11
RESULTS .....	14
Figure 2. ....	15
Figure 3. ....	17
Figure 4. ....	18
Figure 5. ....	22
Figure 6. ....	23
Figure 7. ....	25
Figure 8. ....	27
Figure 9. ....	28
Figure 10. ....	29
Figure 11. ....	30
Figure 12. ....	31
Figure 13. ....	32
Figure 14. ....	33
Figure 15. ....	34
Figure 16. ....	35
Figure 17. ....	36
DISCUSSION .....	37
AUTHORS' CONCLUSIONS .....	39
ACKNOWLEDGEMENTS .....	39
REFERENCES .....	40
CHARACTERISTICS OF STUDIES .....	50
DATA .....	207
Test 1. All .....	208
Test 2. ICDAS .....	209
Test 3. ERK .....	210
Test 4. Other visual .....	210
ADDITIONAL TABLES .....	210
APPENDICES .....	220
WHAT'S NEW .....	220
HISTORY .....	220
CONTRIBUTIONS OF AUTHORS .....	221
DECLARATIONS OF INTEREST .....	221
SOURCES OF SUPPORT .....	221
DIFFERENCES BETWEEN PROTOCOL AND REVIEW .....	221
INDEX TERMS .....	221

[Diagnostic Test Accuracy Review]

# Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries

Richard Macey<sup>1</sup>, Tanya Walsh<sup>1</sup>, Philip Riley<sup>2</sup>, Anne-Marie Glenny<sup>1</sup>, Helen V Worthington<sup>2</sup>, Lucy O'Malley<sup>1</sup>, Janet E Clarkson<sup>3</sup>, David Ricketts<sup>4</sup>

<sup>1</sup>Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. <sup>2</sup>Cochrane Oral Health, Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. <sup>3</sup>Division of Oral Health Sciences, School of Dentistry, University of Dundee, Dundee, UK. <sup>4</sup>School of Dentistry, University of Dundee, Dundee, UK

**Contact:** Tanya Walsh, [tanya.walsh@manchester.ac.uk](mailto:tanya.walsh@manchester.ac.uk).

**Editorial group:** Cochrane Oral Health Group.

**Publication status and date:** Edited (no change to conclusions), published in Issue 12, 2021.

**Citation:** Macey R, Walsh T, Riley P, Glenny A-M, Worthington HV, O'Malley L, Clarkson JE, Ricketts D. Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries. *Cochrane Database of Systematic Reviews* 2021, Issue 6. Art. No.: CD014546. DOI: [10.1002/14651858.CD014546](https://doi.org/10.1002/14651858.CD014546).

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

## ABSTRACT

### Background

The detection and diagnosis of caries at the initial (non-cavitated) and moderate (enamel) levels of severity is fundamental to achieving and maintaining good oral health and prevention of oral diseases. An increasing array of methods of early caries detection have been proposed that could potentially support traditional methods of detection and diagnosis. Earlier identification of disease could afford patients the opportunity of less invasive treatment with less destruction of tooth tissue, reduce the need for treatment with aerosol-generating procedures, and potentially result in a reduced cost of care to the patient and to healthcare services.

### Objectives

To determine the diagnostic accuracy of different visual classification systems for the detection and diagnosis of non-cavitated coronal dental caries for different purposes (detection and diagnosis) and in different populations (children or adults).

### Search methods

Cochrane Oral Health's Information Specialist undertook a search of the following databases: MEDLINE Ovid (1946 to 30 April 2020); Embase Ovid (1980 to 30 April 2020); US National Institutes of Health Ongoing Trials Register (ClinicalTrials.gov, to 30 April 2020); and the World Health Organization International Clinical Trials Registry Platform (to 30 April 2020). We studied reference lists as well as published systematic review articles.

### Selection criteria

We included diagnostic accuracy study designs that compared a visual classification system (index test) with a reference standard (histology, excavation, radiographs). This included cross-sectional studies that evaluated the diagnostic accuracy of single index tests and studies that directly compared two or more index tests. Studies reporting at both the patient or tooth surface level were included. In vitro and in vivo studies were considered. Studies that explicitly recruited participants with caries into dentine or frank cavitation were excluded. 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

We extracted data independently and in duplicate using a standardised data extraction and quality assessment form based on QUADAS-2 specific to the review context. Estimates of diagnostic accuracy were determined using the bivariate hierarchical method to produce summary points of sensitivity and specificity with 95% confidence intervals (CIs) and regions, and 95% prediction regions. The comparative accuracy of different classification systems was conducted based on indirect comparisons. Potential sources of heterogeneity were pre-specified and explored visually and more formally through meta-regression.

## Main results

We included 71 datasets from 67 studies (48 completed in vitro) reporting a total of 19,590 tooth sites/surfaces. The most frequently reported classification systems were the International Caries Detection and Assessment System (ICDAS) (36 studies) and Ekstrand-Ricketts-Kidd (ERK) (15 studies). In reporting the results, no distinction was made between detection and diagnosis. Only two studies were at low risk of bias across all four domains, and 15 studies were at low concern for applicability across all three domains. The patient selection domain had the highest proportion of high risk of bias studies (49 studies). Four studies were assessed at high risk of bias for the index test domain, nine for the reference standard domain, and seven for the flow and timing domain. Due to the high number of studies on extracted teeth concerns regarding applicability were high for the patient selection and index test domains (49 and 46 studies respectively).

Studies were synthesised using a hierarchical bivariate method for meta-analysis. There was substantial variability in the results of the individual studies: sensitivities ranged from 0.16 to 1.00 and specificities from 0 to 1.00. For all visual classification systems the estimated summary sensitivity and specificity point was 0.86 (95% CI 0.80 to 0.90) and 0.77 (95% CI 0.72 to 0.82) respectively, diagnostic odds ratio (DOR) 20.38 (95% CI 14.33 to 28.98). In a cohort of 1000 tooth surfaces with 28% prevalence of enamel caries, this would result in 40 being classified as disease free when enamel caries was truly present (false negatives), and 163 being classified as diseased in the absence of enamel caries (false positives). The addition of test type to the model did not result in any meaningful difference to the sensitivity or specificity estimates ( $\text{Chi}^2(4) = 3.78, P = 0.44$ ), nor did the addition of primary or permanent dentition ( $\text{Chi}^2(2) = 0.90, P = 0.64$ ). The variability of results could not be explained by tooth surface (occlusal or approximal), prevalence of dentinal caries in the sample, nor reference standard. Only one study intentionally included restored teeth in its sample and no studies reported the inclusion of sealants.

We rated the certainty of the evidence as low, and downgraded two levels in total for risk of bias due to limitations in the design and conduct of the included studies, indirectness arising from the in vitro studies, and inconsistency of results.

## Authors' conclusions

Whilst the confidence intervals for the summary points of the different visual classification systems indicated reasonable performance, they do not reflect the confidence that one can have in the accuracy of assessment using these systems due to the considerable unexplained heterogeneity evident across the studies. The prediction regions in which the sensitivity and specificity of a future study should lie are very broad, an important consideration when interpreting the results of this review. Should treatment be provided as a consequence of a false-positive result then this would be non-invasive, typically the application of fluoride varnish where it was not required, with low potential for an adverse event but healthcare resource and finance costs.

Despite the robust methodology applied in this comprehensive review, the results should be interpreted with some caution due to shortcomings in the design and execution of many of the included studies. Studies to determine the diagnostic accuracy of methods to detect and diagnose caries in situ are particularly challenging. Wherever possible future studies should be carried out in a clinical setting, to provide a realistic assessment of performance within the oral cavity with the challenges of plaque, tooth staining, and restorations, and consider methods to minimise bias arising from the use of imperfect reference standards in clinical studies.

## PLAIN LANGUAGE SUMMARY

### Visual or visual-tactile examination for the diagnosis of dental caries

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

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 minimise the number of false-positive results when treatment may be given when caries is absent, and improved visual detection methods may reduce such occurrences.

#### What is the aim of this review?

The aim of this Cochrane Review was to find out how accurate visual classification systems are for detecting early tooth decay as part of the dental 'check-up' for children and adults who visit their dentist. Researchers in Cochrane included 67 studies to answer this question.

#### What was studied in the review?

Two main visual classification systems were studied in this review: the International Caries Detection and Assessment System (ICDAS) and the Ekstrand-Ricketts-Kidd (ERK) system. A third group of visual classifications is reported and labelled as 'Other' because the studies did

not report what system was used. We studied decay on the occlusal surfaces (biting surfaces of the back teeth), the proximal surfaces (tooth surfaces that are next to each other), and smooth surfaces.

### **What are the main results of the review?**

The review included 67 studies with a total of 19,590 teeth. Some studies reported on more than one type of classification system, this gave us 71 sets of data to use. The results of these studies indicate that, in theory, if the visual classification systems were to be used by a dentist for a routine dental examination in a group of 1000 tooth sites/surfaces, of whom 350 (28%) have early tooth decay:

- the use of a visual classification system will indicate that an estimated 403 will have early tooth decay, and of these, 163 (40%) will not have tooth decay (false positive - incorrect diagnosis);
- of the 597 tooth sites/surfaces with a result indicating that tooth decay is not present, 40 (7%) will have early tooth decay (false negative - incorrect diagnosis).

A diagram of these results can be found at [oralhealth.cochrane.org/visual-examination-classification-systems-results-0331c](http://oralhealth.cochrane.org/visual-examination-classification-systems-results-0331c). In this example, visual classification systems produce a high proportion of false-positive results. Treatment in the absence of disease is likely to be non-invasive such as the application of high fluoride toothpaste, or oral health advice and guidance from the dentist, but will incur financial cost to the patient or healthcare provider.

We found no evidence from the data collected that the classification systems differed in their accuracy.

### **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. This is because teeth with deep tooth decay would be easier to identify. However, there were some problems with how the studies were conducted. This may result in the visual classification systems appearing more accurate than they really are, increasing the number of correct visual classification results. We judged the certainty of the evidence to be low due to how the studies selected their participants, the large number of studies that were carried out in a laboratory setting on extracted teeth, and variation in results.

### **Who do the results of this review apply to?**

Studies included in the review were carried out in Brazil, Europe, Japan, and Australia. A large number of studies performed the tests on extracted teeth, while clinical studies were completed in dental hospitals, general dental practices, or schools. Studies were from the years 1988 to 2019.

### **What are the implications of this review?**

We observed substantial variation in the results, which is perhaps unsurprising as the use of these classification systems involve interpretation by the user. There is considerable uncertainty in the likely performance of a future study. Further research studies should be carried out in a clinical setting.

### **How up-to-date is this review?**

Review authors searched for and used studies published up to 30 April 2020.

## SUMMARY OF FINDINGS

### Summary of findings 1. Summary of findings table - main results

Question	What is the diagnostic accuracy of visual classification for the detection and diagnosis of dental caries?	
<b>Population</b>	Asymptomatic children or adults presenting for clinical examination (clinical studies); extracted teeth of children or adults (in vitro studies). Clinical or in vitro studies which intentionally included dentine and frank cavitations for assessment were excluded	
<b>Index test</b>	Visual classification. Results of the index tests were usually recorded on an ordinal scale to indicate severity of disease. For the purposes of this review the positivity threshold was caries in enamel	
<b>Comparator test</b>	Estimates were compared across different classification systems. A separate review in this series explores the comparative accuracy of visual classification, fluorescence-based, radiograph, and transillumination methods of detection and diagnosis	
<b>Target condition</b>	Early dental caries (positivity threshold of early caries or beyond)	
<b>Reference standard</b>	Histology, excavation, radiographs	
<b>Action</b>	If dental caries can be detected at an early stage then remedial action can be taken to arrest, or even reverse decay, and potentially prevent restorations	
<b>Diagnostic stage</b>	Aimed at the general dental practitioner assessing regularly attending patients for early stage caries	
<b>Quantity of evidence</b>	67 studies providing 71 datasets. 6827 lesions in 19,590 tooth surfaces (35% prevalence)	
<b>Findings</b>		
<b>Sensitivity (95% CI)</b>	0.86 (0.80 to 0.90)	There was substantial variability in the results of the individual studies, with sensitivities that ranged from 0.16 to 1.00 and specificities from 0 to 1.00. The relatively narrow confidence intervals and confidence regions for the summary estimates are reflective of the volume of data in the analysis. The prediction region, which gives an indication of between-study heterogeneity and the region where the true sensitivity and specificity of a future study should lie, is very broad
<b>Specificity (95% CI)</b>	0.77 (0.72 to 0.82)	
<b>DOR (95% CI)</b>	20.38 (14.33 to 28.98)	
	The observed heterogeneity could not be explained by tooth surface, prevalence of dentine caries in the study samples, or choice of reference standard	
<b>Outcome</b>	<b>Numbers applied to a hypothetical cohort of 1000 tooth sites or surfaces</b> <b>Effect per 1000 tooth sites or surfaces assessed (95% CI)</b>	<b>Test accuracy</b> <b>Certainty of the evidence</b>
	<b>Pre-test probability 28%<sup>a</sup></b>	

<b>True positives (tooth sites or surfaces with early enamel caries)</b>	240 (225 to 251)	⊕⊕⊕⊕ LOW
<b>False negatives (tooth sites or surfaces incorrectly classified as not having early enamel caries)</b>	40 (29 to 54)	
<b>True negatives (tooth sites or surfaces without early enamel caries)</b>	557 (516 to 592)	⊕⊕⊕⊕ LOW
<b>False positives (tooth sites or surfaces incorrectly classified as having early enamel caries)<sup>b</sup></b>	163 (128 to 204)	
<b>Limitations - factors that may decrease the certainty of the evidence</b>		
<b>Risk of bias</b>	Only 2 studies were considered to be at low risk of bias overall. Across the 4 domains, the patient selection domain had the highest number of studies judged at high risk of bias (49 studies); we judged the index test, reference standard, and flow and timing domains to have a low proportion of studies at high risk of bias (4, 9, and 7 studies respectively). All studies avoided a case-control design, the majority avoiding inappropriate exclusions (96%), but consecutive or random sampling was used in only 10% of studies. Most included studies were in vitro studies using histology as the reference standard, and likely to correctly classify the target condition, however some studies used an imperfect reference standard such as excavation, radiographs, or a visual examination after separation of the approximal surfaces	
<b>Applicability of evidence to question</b>	15 studies were considered to have low concern for applicability across all domains. We judged 49 studies of extracted teeth to be of high concern for applicability of the sample population, with the remaining clinical studies directly applicable to our review question. Concern regarding applicability for the index test domain was also high (46 studies). Most studies were in vitro studies (48 studies), 18 studies carried out the index test evaluations in a clinical setting, 1 study carried out the clinical assessment in a school setting. The dominance of in vitro studies also means that information on how the results of these devices are used to support diagnosis, as opposed to pure detection, is limited	
<b>Test accuracy Certainty of the evidence</b>	We rated the certainty of the evidence as low, and downgraded 2 levels in total, for risk of bias due to limitations in the design and conduct of the included studies, indirectness arising from the in vitro studies, and inconsistency of results	

<sup>a</sup>Hypothetical cohorts of 1000 lesions are presented for numbers estimated at 28% enamel caries prevalence taken from the level of cavitated teeth in the UK Adult Dental Health Survey (Steele 2011). Based on consultation with clinical colleagues this illustrative prevalence value addresses concern that the overall prevalence observed in the data (35%) is not representative of the general population.

<sup>b</sup>Despite the relatively high proportion of false positives, should care be given where it was not required early, non-invasive treatment would likely be limited to the application of fluoride varnish, which has a cost implication, but there would be a low potential for an adverse event for the patient.

CI: confidence interval; DOR: diagnostic odds ratio.

## Summary of findings 2. Summary of findings table - comparison of tests and dentitions

Question	<b>What is the diagnostic accuracy of visual classification for the detection and diagnosis of early dental caries?</b>
Population	Asymptomatic children or adults presenting for clinical examination (clinical studies); extracted teeth of children or adults (in vitro studies). Clinical or in vitro studies which intentionally included dentine and frank cavitations for assessment were excluded
Index test	Visual classification. Results of the index tests were usually recorded on an ordinal scale to indicate severity of disease. For the purposes of this review the positivity threshold was caries in enamel
Comparator test	Estimates were compared across different classification systems. A separate review in this series explores the comparative accuracy of visual classification, fluorescence-based, radiograph, and transillumination methods of detection and diagnosis
Target condition	Early dental caries
Reference standard	Histology, excavation, radiograph
Action	If dental caries can be detected at an early stage then remedial action can be taken to arrest or even reverse decay, and potentially prevent restorations
Diagnostic stage	Aimed at the general dental practitioner assessing regularly attending patients for early stage caries
Quantity of evidence	67 studies providing 71 datasets. 6827 enamel lesions in 19,590 tooth surfaces (35% prevalence)

### Findings: analysis comparing ERK, ICDAS, and 'other visual' tests

Test	Datasets	Tooth surfaces	Sensitivity (95% CI)	Specificity (95% CI)	Interpretation
<b>ERK</b>	15	1306	0.85 (0.74 to 0.92)	0.72 (0.61 to 0.81)	These results should be interpreted taking into account the factors that limit the certainty of the evidence as indicated in <a href="#">Summary of findings 1</a> . The similarity of the summary sensitivity and specificity points across classification systems does not necessarily mean that they are similarly accurate, only that the middle of the observed data is similar between classification systems
<b>ICDAS</b>	38	6817	0.88 (0.81 to 0.92)	0.76 (0.67 to 0.83)	
<b>Other</b>	18	11,467	0.81 (0.65 to 0.91)	0.84 (0.73 to 0.91)	

No overall difference in test sensitivity and specificity:  $\text{Chi}^2(4) = 3.78, P = 0.44$

Using ICDAS as the reference category:  
 difference in sensitivity for ERK 0.02 (-0.08 to 0.13),  $P = 0.68$ ; other visual 0.07 (-0.07 to 0.21),  $P = 0.36$





difference in specificity for ERK 0.04 (-0.09 to 0.16), P = 0.54; other visual -0.08 (-0.20 to 0.03), P = 0.16

**Findings: analysis comparing dentition. Consequences in a cohort of 1000 tooth surfaces**

Dentition	Datasets	Tooth surfaces	Sensitivity (95% CI)	Specificity (95% CI)	Interpretation
Primary	29	3705	0.87 (0.80 to 0.92)	0.79 (0.70 to 0.85)	These results should be interpreted taking into account the factors that limit the certainty of the evidence as indicated in <a href="#">Summary of findings 1</a> . The similarity of the summary sensitivity and specificity points across dentitions does not necessarily mean that they are similarly accurate, only that the middle of the observed data is similar
Permanent	42	15,885	0.85 (0.77 to 0.90)	0.76 (0.69 to 0.83)	
No overall difference in test sensitivity and specificity: Chi <sup>2</sup> (2) = 0.90, P = 0.64			Difference 0.02 (95% CI -0.07 to 0.11) P = 0.61	Difference 0.02 (95% CI -0.08 to 0.13) P = 0.69	

CI: confidence interval; df: degrees of freedom; ERK: Ekstrand-Ricketts-Kidd system; ICDAS: International Caries Detection and Assessment System.

## BACKGROUND

Cochrane Oral Health (COH) has undertaken several Cochrane Reviews of diagnostic test accuracy (DTA) on the detection and diagnosis of dental caries (Fee 2020; Macey 2020; Macey 2021; Macey 2021a; Walsh 2021). The suite of systematic reviews forms part of a UK National Institute for Health Research (NIHR) Cochrane Programme Grant and involved collaboration with the Complex Reviews Support Unit. The reviews follow standard Cochrane DTA methodology and are 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, preventive oral health programmes have almost eradicated caries, but such activities have not been widely replicated in other locations (Pitts 2017). The 2015 Global Burden of Disease study has 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, the global incidence of untreated caries was reported to be 2.4 billion in 2010 (Feigin 2016; Kassebaum 2015; World Health Organization 2017) and despite a reduction in caries in some industrialised countries, the global incidence of caries has increased by 14% in the five years to 2015 to over half a billion people (Feigin 2016). In the UK, recent statistics indicate that the primary reason for childhood (aged 5 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, and that the disease trajectory will be steeper than those without early caries experience (Broadbent 2008; Hall-Scullin 2017).

Untreated caries can lead to episodes of severe pain and infection, often requiring treatment 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 2009b; 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). Caries diagnosis is a complex process involving the staging of the disease process (severity), assessment of the location of the lesion (whether in a plaque stagnation area or not), understanding caries risk factors for the individual patient and lesion and assessment of lesion activity, but pivotal to the diagnostic process is the early identification of caries, namely caries detection, which is the focus of this suite of reviews.

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). Dental caries if left unchecked is likely to progress in severity based upon the amount of demineralisation that has taken place within the lesion and the depth of penetration of the lesion into the tooth tissues. These two factors are not necessarily linearly related but both will dictate whether a lesion develops a frank hole or discontinuity of the tooth surface (cavitation). As such, a lesion may be confined to enamel in depth and may (rarely) or may not be cavitated based upon the amount of demineralisation and similarly the same argument can apply to lesions extending into dentine histopathologically. Thus lesion depth estimation and surface topography are important features to capture in caries detection. The prevalence of this early caries state is not often reported in dental epidemiology, most reports preferring to focus on cavitated/dentinal lesions which may be easier to detect. For example, the most recent UK Adult Dental Health survey reported that 31% of the sample had untreated caries into dentine (Steele 2011; White 2012), and a US study reported levels of cavities at 15.30% 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 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 softer demineralised tissue with resins for initial caries;
- operative interventions such as step-wise caries removal and restoration for extensive lesions.

With advances in technology over the last two decades, alternative methods of caries detection have become available, such as advancements in radiography and the development of fluorescence, transillumination, 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. Furthermore, the ability to accurately detect early caries and prevent early caries lesions from progressing has become increasingly important in reducing the

need to undertake/undergo invasive treatment which may require the use of aerosol generating procedures (AGPs).

### 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 due to bacterial breakdown of sugars in the diet which leads to production of acid and demineralisation of the tooth. 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 D<sub>1</sub>, D<sub>2</sub>, D<sub>3</sub>, and D<sub>4</sub> (Anaise 1984) (Additional Table 1). Typically the D<sub>3</sub> threshold has been used to determine the presence of caries (Pitts 1988; Shoab 2009).

These four categories have formed the basis for expanded indices such as the International Caries Detection and Assessment System (ICDAS) (Ekstrand 2007; Ismail 2007). Other available systems include: the Nyvad system (Nyvad 1999); Ekstrand-Ricketts-Kidd (ERK) system (Ekstrand 1997); British Association for the Study of Community Dentistry (BASCD) (Pitts 1997); and the Dundee Selectable Threshold Method for caries diagnosis (DSTM) (Fyffe 2000a).

### 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 softer 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:

1. when dentition is sound the clinician proceeds with preventative strategies to prevent sound surfaces from developing caries;
2. non-invasive treatment of the lesion to arrest the decay process and encourage remineralisation, preventing initial lesions from progressing to cavitated decay; and
3. management of more severe caries through excavation and restoration or potentially extraction.

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-tactile clinical examination, and the ability of clinicians to accurately detect disease in this way has been researched for over half a century (Backer Dirks 1951). Tests may be suitable at different stages of the care pathway (Bloemendal 2004; Fyffe 2000a), and the use of additional detection tools can add to the detection, diagnosis, and monitoring process. The generic protocol (Macey 2018) provides information regarding alternative index tests for caries detection and diagnosis in this suite of Cochrane Reviews.

This review focuses on the conventional and enhanced visual or visual-tactile examinations, and incorporates various approaches to the visual examination using different classifications of disease; these include ICDAS, ERK, and Nyvad (Additional Table 2).

### Clinical pathway

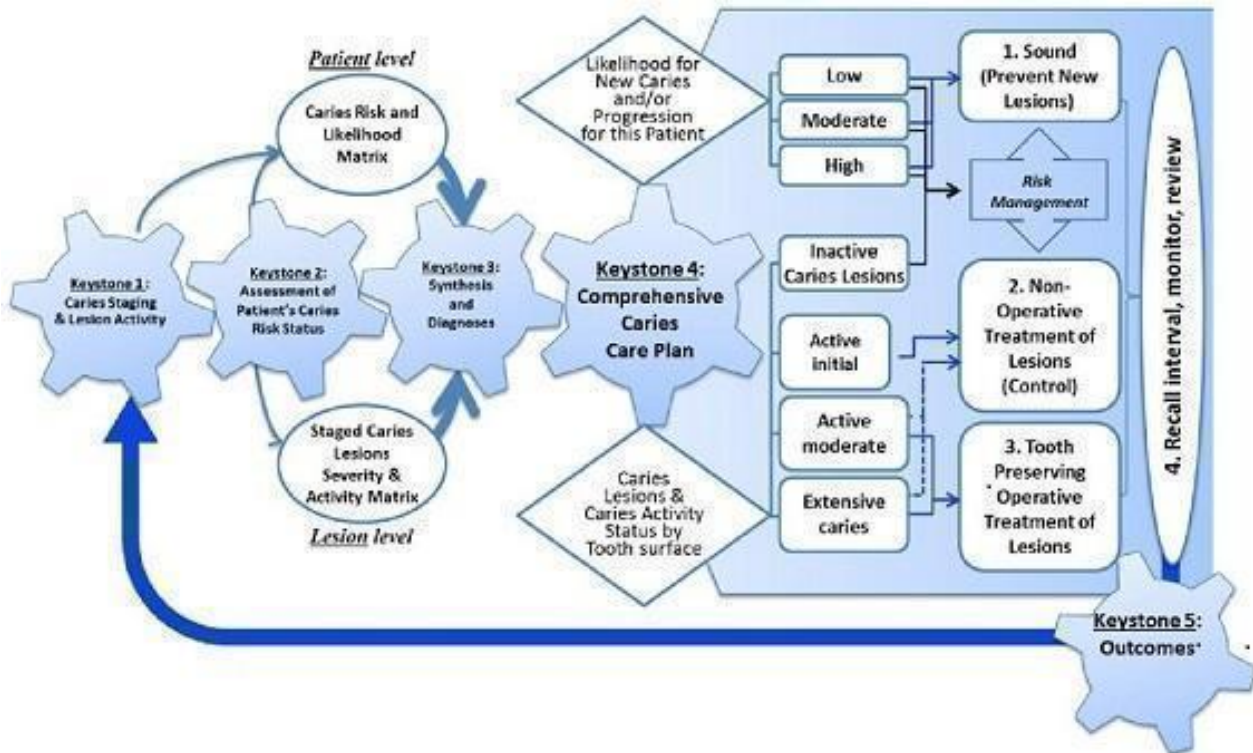
The process proceeding from a patient attending for a routine dental examination and a caries assessment being undertaken has four intertwined stages: screening, detection, diagnosis, and treatment planning. If the presenting patient is seemingly asymptomatic then this could be viewed as a screening exercise, as the clinician is seeking to establish the presence or absence of disease (Wilson 1968). However, patients are likely to present with some degree of caries as the established classification systems (for example ICDAS) are sensitive enough to detect any changes in the enamel of the tooth's surface as confirmed by a survey of the English population reporting 0.8 carious teeth on average (Adult Dental Health Survey 2009). Therefore, detection is a more reasonable description of this initial examination, this is where the clinician aims to establish the true presence or absence of disease. Since caries is a dynamic process the pure detection of the disease at one time point is not sufficient to inform the future care of the patient, additionally the depth and severity of demineralisation, allied to a decision on 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 care, 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). ICCMS has been developed further for use in primary care namely Caries Care International (Martignon 2019).

Figure 1 presents the key elements of the ICCMS process and this review 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)**

In clinical practice, a conventional visual or visual-tactile oral examination would always be undertaken as part of the clinical examination. The methods used in the visual examinations evaluated in this systematic review may be different to those currently used by general dental practitioners in routine examinations. For example, a full ICDAS examination and charting can take up to 20 minutes to complete, which may be considered impractical under the time pressures associated with the routine dental examination. The information from caries detection (including assessment of severity of disease) will be an integral part of diagnosis, which additionally incorporates patient history, risk factors, and treatment planning protocols.

**Alternative test(s)**

Other reviews have been completed as part of this series of Cochrane Reviews and report the alternative tests that are available to the dental clinician, they include.

- Fluorescence (Macey 2020): 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).

- Radiography (Walsh 2021): bitewing radiology is the most commonly used method. Other techniques include subtraction radiography which produces a semi-automated method for monitoring progression of lesions (Ellwood 1997; Wenzel 2006) 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).
- Transillumination (Macey 2021a): fibre optic transillumination (FOTI) which 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.
- Electrical conductance (Macey 2021): 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).

For more details please see the protocol for this review (Macey 2018).

## Rationale

Despite technological advancement, the typical method of caries detection is currently based upon information from visual-tactile clinical examination, supplemented with radiographs where indicated. There have been a number of systematic reviews of visual or visual-tactile examinations. Bader 2002 completed an extensive review of in vitro studies investigating visual, radiographic, fibre optic, electrical conductance, and fluorescence in the primary and permanent dentition. This review was limited to studies with a histological reference standard and grouped studies according to index test, disease threshold (enamel or dentinal lesions), and tooth surfaces (occlusal or proximal); a meta-analysis was not undertaken and the authors graded the quality of the available evidence as low (Bader 2002). This review predates the development of meta-analysis methods for DTA reviews recommended in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (Deeks 2013). More recently a review has been completed which investigated visual techniques, which included primary and permanent dentition, occlusal and proximal surfaces, and accepted histological, operative, visual, and radiographs as the reference standard (Gimenez 2015) however, they did not use hierarchical methods in their meta-analysis.

In this Cochrane Review we have included contemporary studies irrespective of publication language and status, and built upon existing research to incorporate methodological developments by: expanding the search strategy to capture all relevant evidence, applying appropriate hierarchical analysis (Dinnes 2016), and assessing the body of evidence using GRADE (Schünemann 2020) 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 visual classification systems for the detection and diagnosis of non-cavitated coronal dental caries in children and adults.

The specific research questions addressed in this systematic review were.

- What is the diagnostic test accuracy of different methods of visual classification for:
  - different purposes (detection or diagnosis);
  - in different populations (children: primary/mixed dentition, adolescents: immature permanent dentition, or adults: mature permanent dentition);
  - when a comparison is made between different visual tests.

## Secondary objectives

We investigated the following areas of potential heterogeneity:

- the use of different reference standards;
- tooth surface (occlusal, proximal, or smooth surface);

- prevalence of dentinal caries;
- participants or teeth with previously applied restorations (secondary caries) and pit and fissure sealants.

## 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 compared 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 novel systems may be identifiable through this design alone;
- reporting at either the patient, tooth, or tooth surface level, however only those reporting at the tooth surface level were included in the primary analysis;
- in vivo or in vitro studies;
- prospective studies that evaluated the diagnostic accuracy of single index tests, studies that directly compared two or more index tests of this type and studies that attempted to detect caries on surfaces with fissure sealants.

In vitro studies are those in which teeth had been extracted prior to the initiation of the study, and the index test and a histological reference standard were applied. This scenario is not representative of the typical clinical setting. In vivo studies recruited participants and conducted the index tests and reference standards with the teeth in the oral cavity, usually without extraction of the teeth. For in vivo studies histology would not usually be undertaken with the exception of teeth indicated for extraction or primary teeth close to exfoliation.

Studies were ineligible for inclusion where:

- artificially created carious lesions were used in the assessment;
- studies used a visual test as part of a combined test with an adjunct such as radiography or fluorescence (these studies were included in the systematic reviews of this series);
- an index test was used during the excavation of dental caries to ascertain the optimum depth of excavation.

#### Participants

Participants seemingly asymptomatic for dental caries but may have early caries which is 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

Visual or visual-tactile examination on intact teeth according to detailed criteria and indices (e.g. Ekstrand, International Caries Detection and Assessment System (ICDAS)) (Additional Table 2). The tactile label infers the use of a sharp- or ball-ended probe which may additionally be used in some studies to detect for example stickiness in pits and fissures or discontinuity in the tooth surface.

Where studies investigated multiple index tests, it was important to be able to assess the visual index tests in isolation, otherwise the result of one index test may influence another.

Where studies used multiple examiners the most appropriate examiner to the research question was selected. For example, if the study used dental students, general dental practitioners and restorative consultants, then the results of the general dental practitioner were chosen. In the scenario where multiple examiners were stated to have similar skills and experience, then the mean sensitivity and specificity values were extracted, if this was not available then the first set of reported results was selected.

### Target conditions

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

### Reference standards

A number of different reference standards have been used in primary studies for caries detection and diagnosis. The only way to achieve a true diagnosis of caries presence and depth is to extract and section the tooth and then perform a histological assessment (Downer 1975; Kidd 2004). This approach is commonly undertaken on previously extracted teeth for in vitro studies but unethical for a healthy population in clinical (in vivo) studies. The only scenario where histology could be appropriate for studies undertaken in a primary or secondary care dental 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), the index test could be applied prior to extraction, and followed by a histological reference standard.

A clinician may use a combination of approaches including enhanced visual, radiograph and/or fluorescence tests to decide whether further excavation of the tooth is warranted, the caries would then be removed with a dental burr (drill) in preparation for a restoration and the depth of decay reported. This composite set of tests could be used as the reference standard. There is concern regarding the accuracy of radiographs as a reference standard to detect early enamel lesions and this was accounted for in the quality assessment.

Tooth separation using orthodontic bands was deemed an acceptable reference standard for studies investigating approximal surfaces. This method involves placing a band between the mesial and distal surfaces which is typically left in situ for up to a week

to encourage tooth movement and allow a direct view of the approximal surface of the tooth which would otherwise have been obscured by the abutting tooth. Although not often used in general practice, this approach does make in vivo studies of approximal surfaces more achievable and potentially more robust than the excavation or radiograph options, but may result in incorporation bias.

The optimum reference standard was histology. Tooth separation, radiographs, and operative exploration were considered to be acceptable reference standards for in vivo studies.

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

## Search methods for identification of studies

### 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 30 April 2020) (Appendix 1);
- Embase Ovid (1980 to 30 April 2020) (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](http://clinicaltrials.gov); searched 30 April 2020) (Appendix 3);
- World Health Organization International Clinical Trials Registry Platform ([apps.who.int/trialsearch](http://apps.who.int/trialsearch); searched 30 April 2020) (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. Studies were excluded if they failed to present the data in a 2 x 2 contingency table, or failed to report sufficient information to enable a 2 x 2 table to be constructed. In such instances the study authors were contacted and the required data requested. An adapted PRISMA flowchart is used to report the study selection process (McInnes 2018).

Where studies reported combinations of different index tests the results were not included in this review unless the diagnostic information could be isolated to ensure the results were for visual tests alone. These combined tests (e.g. visual and radiograph combined) were included in the review of the additional test as a combined test result (e.g. the radiograph review (Walsh 2021)) and were noted in the reasons for exclusion of this review.

## Data extraction and management

Once agreement for inclusion was reached, the studies were categorised according to their index test - specifically the type of visual classification that the examiners used, the tooth surface, and the dentition of the participants. Two review authors extracted data independently and in duplicate using a piloted data extraction form based on the review inclusion criteria. Disagreements were resolved through discussion by the review team. Where data had been reported for multiple 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, 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 the 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](#)). Two review authors independently completed QUADAS-2. Any disagreements were resolved through discussion and, where necessary, consultation with another member of the author team.

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 would be 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 quality assessment were presented numerically and graphically.

### Participant selection domain

The selection of patients has a fundamental effect on the estimated accuracy of an index test. The disease stages of sound and carious enamel should be represented in the sample and children, adolescents, and adults should be represented in the included studies to allow a complete appraisal of a test's potential to correctly classify disease in different populations.

It was acceptable for studies to focus on early enamel lesions for a specific surface (occlusal, approximal, or smooth) or dentition (primary or mixed, immature permanent, permanent). For a low

risk of bias judgement the inclusion of study participants or teeth meeting the eligibility criteria should be consecutive or random as inappropriate exclusions may lead to an over- or under-estimation of the test's ability to detect disease. Additionally the prevalence and severity of disease reported 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, due to the potential for the index test and reference standard to be reporting on different sites within the same mouth. Case-control studies were considered to be at a high risk of bias.

### Index test domain

The nature of the index tests and the visual presentation of the target condition means that it is feasible to ensure that the index test is conducted prior to the reference standard. The visual 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. The threshold of disease positive and negative should have been determined prior to analysis and be reflective of the participants recruited to the study.

Where both initial and more severe assessments of disease are considered within a primary study, given the subjective nature of the visual classifications there may be potential for information bias unless different examiners have been undertaken assessments for each of the different thresholds assessed within a study. For example, if the assessor's judgement is uncertain between caries into enamel or caries into dentine, the interpretation of the first threshold would influence the decision made on the second threshold.

### Reference standard domain

To minimise potential for bias, it is preferable for separate examiners to carry out the index test and reference standard. If the reference standard was a radiograph, excavation, or tooth separation to allow a visual assessment of the approximal surface then it should be undertaken by a different examiner from that completing the index test, as the subjectivity of these reference standards could be compromised by knowledge of the index test results. An exception for this signalling question was applied where the tooth had been extracted, sectioned, and prepared for histological evaluation, as it is extremely unlikely that the examiner would be able to recall the specific tooth or participant and associated index test result.

Ideally, each participating tooth or patient within a study should receive the same reference standard. This is possible in the in vitro setting where a histological assessment can be applied to each tooth. If a study allocated participants or specific teeth to different reference tests then reasons for this allocation should have been clearly reported. This could occur where teeth are excavated, as only those teeth that are considered to be cavitated could be justifiably excavated, so a prior reference standard would have to be applied which would either be a radiograph or the visual assessment. This was considered to represent a high risk of bias due to contamination of the reference standard results and the inability to have confidence in the accuracy of these findings.

### Flow and timing domain

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.

### Statistical analysis and data synthesis

In accordance with the focus of this review we stipulated a positivity threshold of enamel caries. Estimates of diagnostic accuracy were expressed as sensitivity and specificity with 95% confidence intervals (CI) for each study and for each available data point if there were multiple index tests, dentition, or surfaces reported within a single study. When there were two or more test results reported in the same study, different dentitions or tooth surfaces evaluated for example, we included them as separate datasets. We also illustrated the 95% confidence and prediction regions to support interpretation. The prediction region indicates where 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.

Hierarchical models were used for data synthesis. The data were extracted for the target condition of early caries (caries into dental enamel) at the tooth surface level. This target condition has been consistently used across the suite of caries detection reviews. Study estimates of sensitivity and specificity were plotted on coupled forest plots and in receiver operating characteristic (ROC) space. A meta-analysis was conducted which combined the results of studies for each index test using a bivariate model to estimate the summary values of sensitivity and specificity at a common threshold (Chu 2006; Reitsma 2005). Data were input to Review Manager 5 (Review Manager 2020) and displayed in the coupled forest plots. Analysis was conducted using xtmelogit and the METANDI package in Stata (Harbord 2009; Stata 14; Takwoingi 2016), and the MetaDTA interactive web-based tool (Freeman 2019). We used meta-regression with xtmelogit to compare the accuracy of different classification systems and different dentitions in this review. We added the classification system and dentition as covariates to the bivariate model, assuming equal variances for the different levels of the covariates, and used a likelihood ratio test to formally assess the significance of any model comparisons (Macaskill 2010; Takwoingi 2016). Initially we allowed the covariate effects 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 (Takwoingi 2016). Where sufficient studies were available we did not assume equal variances of the covariates and modelled separate variances.

### Investigations of heterogeneity

The investigation of each potential source of heterogeneity was considered individually. Initially, a visual 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. Where sufficient numbers of studies allowed, meta-regression analyses were carried out to explore possible sources of heterogeneity. Formal model comparisons were undertaken as previously indicated.

The sources of heterogeneity (specified a priori) were different reference standards used in in vitro and in vivo studies, tooth surface, prevalence of caries into dentine, and studies including previously applied restorations (secondary caries) or pit and fissure sealants. Each potential source of heterogeneity was investigated separately.

### Sensitivity analyses

Where a sufficient number of studies investigated the same index test, the following sensitivity analyses were performed by removing studies from the meta-analysis. This enabled us to assess the impact on summary estimates of restricting the analyses according to studies that meet the following criteria:

- low risk of bias on their inclusion criteria for caries threshold;
- low prevalence of dentine caries (i.e. less than 15%);
- low risk of bias for an index test;
- low risk of bias 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 visual index tests and for the main target conditions following GRADE methods (Hsu 2011; Schünemann 2020; Zhang 2019), and using the GRADEPro online tool ([www.guidelinedevelopment.org](http://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 indirectness of the evidence, the inconsistency of the results, the imprecision of the estimates, and the risk of publication bias. We categorised the certainty of the body of evidence, as high, moderate, low, or very low.

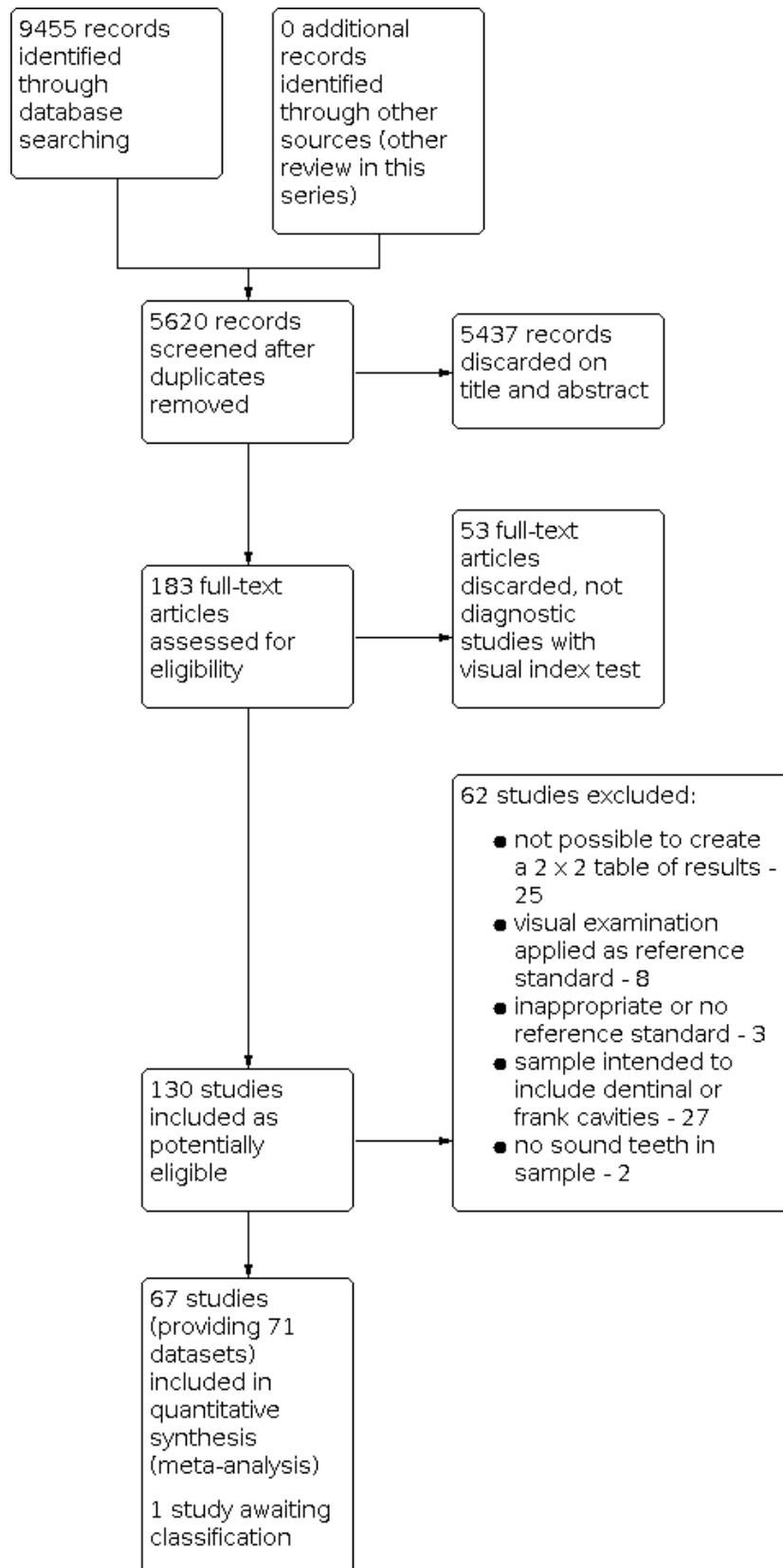
## RESULTS

### Results of the search

The search identified 9455 results, of which 5620 remained after the removal of duplicates. After an initial screening of titles, abstracts, and full texts where necessary, 131 studies were considered to be potentially eligible for inclusion. Upon closer inspection this number reduced to the 67 studies which are included in this review and one study that is awaiting classification (Figure 2). A common reason for exclusion was the inability to create a 2 x 2 table of the results (25 studies). We also excluded 27 studies that detailed their intention to include frankly cavitated teeth or those with dental caries. Studies and their reasons for exclusion are detailed in the [Characteristics of excluded studies](#) table.



**Figure 2. Study flow diagram.**



There are studies included in the analysis twice because they reported results from both the primary and permanent teeth (Ekstrand 2011; Rodrigues 2009), one study reported approximal and occlusal surfaces (Hintze 2003), and one investigated two different scoring systems (Braga 2010). This resulted in 71 included datasets for the analysis, which reported a total of 19,590 sites or surfaces of teeth. 48 of the studies have been included in other reviews of this series on detection of enamel dental caries; of these 25 appeared in more than one of the other reviews in this series (Ashley 1998; Bahrololoomi 2015; Braga 2009; Bussaneli 2015; Bussaneli 2015a; Cinar 2013; Costa 2002; Diniz 2011; Diniz 2012; Diniz 2019; Goel 2009; Jablonski-Momeni 2012; Ko 2015; Kockanat 2017; Kucukyilmaz 2015; Mendes 2006; Neuhaus 2011; Novaes 2009; Novaes 2012; Pereira 2011; Rocha 2003; Rodrigues 2008; Souza 2013; Teo 2014; Tonkaboni 2019), 11 also investigated fluorescence devices (Achilleos 2013; Akarsu 2006; Castilho 2016; Huth 2010; Iranzo-Cortes 2017; Kim 2017; Paula 2011; Rodrigues 2009; Seremidi 2012; Shi 2000; Sridhar 2009), six radiographs (Braun 2017; Da Silva 2010; Ekstrand 2011; Erten 2005; Freitas 2016; Hintze 2003), and six also investigated transillumination devices (Laitala 2017; Nakagawa 2013; Nakajima 2014; Shi 2000; Shimada 2010; Xiao-Hua 2016). All included studies were published between 1988 and 2019, with 45 (67%) studies conducted from 2010 onwards. 23 of the studies originated from Brazil (33%), six from Switzerland, six from Germany, five from each of Turkey and the UK, three from each of Greece and Japan; the remainder came from Asia (eight), other European countries (six), and one from each of Colombia and Australia.

Many different classification systems were used as an index test in the included studies. There are a number of commonly used classification systems that have been regularly used in the literature (Additional Table 2). Most included studies assessed the International Caries Detection and Assessment System (ICDAS) (36 studies and 38 datasets) followed by the Ekstrand-Ricketts-Kidd system (ERK) (15 studies and 15 datasets). Of the remaining 16 studies, two used Nyvad (Braga 2010; Bussaneli 2015), one used the Downer method (Ashley 1998), and one used the Lussi scale (Costa 2002). The remaining studies did not clearly state the classification system used, although one appeared to be very similar to ERK (Haak 2002) and two reported methods which closely resembled ICDAS (Kim 2017; Nakagawa 2013). One study (Kuhnisch 2009a) used a newly devised system called UniViSS, the remainder used procedures that allowed the classification of surfaces as either having sound, enamel, or dentine caries (Forgie 2003; Hintze 2003; Kucukyilmaz 2015; Nakajima 2014; Paula 2011; Rodrigues 2009; Shi 2000; Sidi 1988). Only one study presented a comparative accuracy of visual classification systems (Braga 2010), reporting results for ICDAS and Nyvad.

Occlusal surfaces were investigated in 51 (72%) datasets, approximal surfaces in 18 (25%) datasets and smooth surfaces in two (Nakagawa 2013; Shi 2000). 42 (59%) datasets involved

the permanent dentition and 29 (41%) used the primary or mixed dentition. A reference standard of histology was used in 60 (85%) datasets, of these 49 datasets reported that extracted teeth were used for the index test and reference standard (in vitro studies). In the remaining 11 datasets (in vivo studies) the index test was applied to teeth within the oral cavity that were identified as being due for exfoliation or extraction, and histological assessment was subsequently performed when the teeth were extracted. Studies that did not use histology as a reference standard relied on radiographs (Carvalho 2018 (in vitro study); Kim 2017; Laitala 2017; Sidi 1988), separation of teeth using orthodontic bands to observe approximal surfaces (Bussaneli 2015; Novaes 2009; Novaes 2010) or excavation. Jablonski-Momeni 2012 (in vitro study) performed excavation of all teeth as they were an extracted sample, Akarsu 2006 used a combination of visual, radiograph, and fluorescence tests to decide whether caries was at a severity requiring excavation, and Bahrololoomi 2015 did not provide clarity on how it arrived at the decision to excavate or how those below the threshold were defined. Huth 2010 used a combination of radiographs for all and additionally excavation where caries was suspected. Most studies were in vitro studies (48 studies), 18 studies carried out the index test evaluations in a clinical setting, one study carried out the clinical assessment in a school setting.

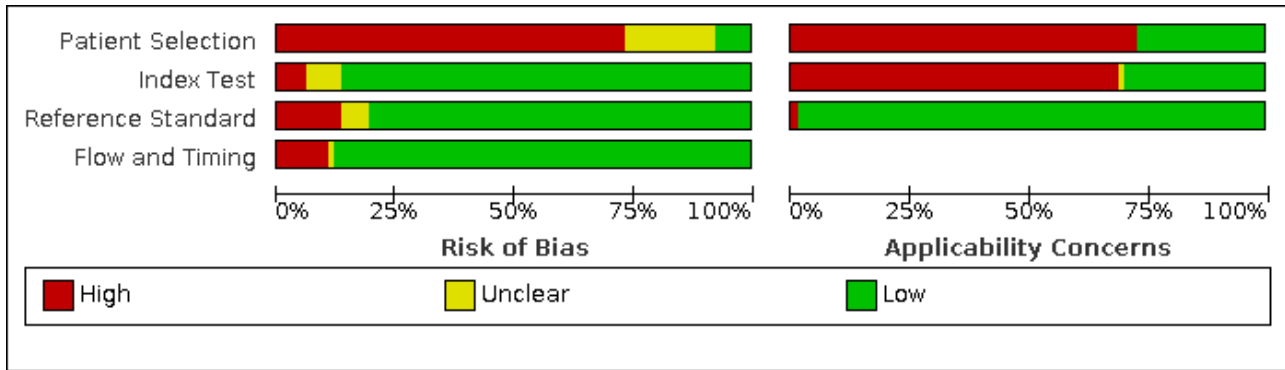
Only one study intentionally included restored teeth in its sample (Kim 2017) and no studies reported the inclusion of sealants.

Key study characteristics are listed in Additional Table 4.

### Methodological quality of included studies

The methodological quality of the 67 included studies is summarised across the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) domains in Figure 3 and the individual study results are shown in Figure 4. There are three study entries that show no assessment of quality, these studies provided multiple datasets: permanent and primary dentition (Ekstrand 2011a; Rodrigues 2009a), and occlusal and approximal surfaces (Hintze 2003a), and the second entry allowed us to include both elements in the meta-analysis, the study quality does not differ from the judgements made in the original entry. Only two studies were considered to be at low risk of bias across all four domains (Braga 2010; Castilho 2016), and 15 studies were considered to be at low concern for applicability across all three domains (Akarsu 2006; Bahrololoomi 2015; Braga 2010; Castilho 2016; Cinar 2013; Diniz 2012; Freitas 2016; Goel 2009; Huth 2010; Kucukyilmaz 2015; Laitala 2017; Novaes 2009; Novaes 2010; Rocha 2003; Teo 2014). Only two studies were judged to be at low risk of bias overall and with low concerns for applicability overall (Braga 2010; Castilho 2016). The patient selection domain had the highest proportion of high risk of bias (49 studies); we judged the index test, reference standard, and flow and timing domains to have a low proportion of studies at high risk of bias (4, 9, and 7 studies respectively).

**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
Achilleos 2013	-	+	+	+	-	-	+
Akarsu 2006	?	+	-	-	+	+	+
Ashley 1998	-	+	+	+	-	-	+
Bahrololoomi 2015	-	-	-	-	+	+	+
Bottenberg 2016	-	+	?	+	-	-	+
Braga 2009	-	?	+	+	-	-	+
Braga 2010	+	+	+	+	+	+	+
Braun 2017	?	+	+	+	-	+	-
Bussaneli 2015	-	+	-	+	-	+	+
Bussaneli 2015a	-	+	?	+	-	-	+
Campos 2017	?	+	+	+	-	-	+
Carvalho 2018	-	+	+	+	-	+	+
Castilho 2016	+	+	+	+	+	+	+
Chawla 2012	-	+	+	+	-	-	+
Cinar 2013	-	+	+	+	+	+	+
Costa 2002	-	+	+	+	-	-	+
Da Silva 2010	-	+	+	+	-	-	+
Diniz 2009	-	+	+	+	-	-	+
Diniz 2011	-	+	+	+	-	-	+
Diniz 2012	?	+	+	+	+	+	+
Diniz 2019	?	+	+	+	-	-	+
Ekstrand 2007	-	+	+	+	-	-	+
Ekstrand 2011	-	+	+	+	-	-	+
Ekstrand 2011a							
Erten 2005	-	+	+	+	-	-	+
Forgie 2003	-	+	+	+	-	-	+

**Figure 4. (Continued)**

Forgie 2003	-	+	+	+	-	-	+
Freitas 2016	-	+	+	?	+	+	+
Goel 2009	-	?	+	+	+	+	+
Haak 2002	?	+	+	+	-	-	+
Hintze 2003	-	+	+	+	-	-	+
Hintze 2003a							
Huth 2010	+	+	-	-	+	+	+
Iranzo-Cortes 2017	-	+	+	+	-	-	+
Iranzo-Cortes 2018	-	+	+	+	-	-	+
Jablonski-Momeni 2008	-	+	+	-	-	-	+
Jablonski-Momeni 2012	?	+	+	+	-	-	+
Kim 2017	?	+	-	-	-	+	+
Ko 2015	-	+	+	+	-	+	+
Kockanat 2017	?	+	+	-	+	-	+
Kucukyilmaz 2015	-	+	+	+	+	+	+
Kuhnisch 2009	-	+	+	+	-	-	+
Kuhnisch 2009a	-	+	+	+	-	-	+
Laitala 2017	?	+	-	+	+	+	+
Mendes 2006	-	?	+	+	-	-	+
Mitropoulos 2010	-	+	+	+	-	-	+
Nakagawa 2013	-	+	+	+	-	-	+
Nakajima 2014	-	+	+	+	-	-	+
Neuhaus 2011	-	-	?	+	-	-	+
Neuhaus 2015	-	+	+	+	-	-	+
Novaes 2009	+	+	-	+	+	+	+
Novaes 2010	+	+	-	+	+	+	+
Novaes 2012	-	?	+	+	-	-	+
Paula 2011	-	+	+	+	-	-	+
Pereira 2011	-	+	+	+	-	-	+
Qudeimat 2019	-	+	+	+	+	-	+
Rocha 2003	-	-	+	+	+	+	+

**Figure 4. (Continued)**

Rocha 2003	+	+	+	+	+	+	+
Rodrigues 2008	-	+	+	+	-	-	+
Rodrigues 2009	-	+	+	+	-	-	+
Rodrigues 2009a							
Rodrigues 2013	?	+	+	+	-	-	+
Seremidi 2012	-	+	+	+	-	-	+
Shi 2000	-	-	+	-	-	-	+
Shimada 2010	-	+	+	+	-	-	+
Sidi 1988	-	?	-	+	-	?	+
Souza 2013	-	+	+	+	-	-	+
Soviero 2012	-	+	+	+	-	-	+
Sridhar 2009	-	+	+	+	+	-	+
Teo 2014	?	+	+	+	+	+	+
Tonkaboni 2019	?	+	?	+	-	-	+
Xiao-Hua 2016	-	+	+	+	-	-	+

- High     
 ? Unclear     
 + Low

**Participant selection**

Seven studies reported the use of random or consecutive sampling (Braga 2010; Castilho 2016; Huth 2010; Laitala 2017; Nakagawa 2013; Novaes 2009; Novaes 2010). Studies failed to apply a random or consecutive sampling technique of the population in 49 studies (73%), instead choosing to select teeth. In 12 studies it was not possible to determine how the sample of participants was obtained and an unclear judgement was made. All studies avoided a case-control and randomised controlled trial (RCT) design. 36 studies avoided inappropriate exclusions (54%), 29 (43%) failed to provide sufficient detail in the description of the sample population to satisfy us that exclusions were appropriate. We identified two studies which attempted to exclude all sound teeth from the sample, this introduced a high risk of bias to the studies (Carvalho 2018; Nakagawa 2013). Both of these studies could be included in this review because although they aimed to exclude sound teeth there were some that were reported in the results. Applicability of the sample population was of high concern for 49 studies (73%), these are the studies which investigated extracted teeth; the objective of this review was to inform general clinical practice and it could not be considered that laboratory (in vitro) studies were clinically relevant. The remaining 18 studies were completed in vivo and of low concern.

**Index test**

Index tests were interpreted without the knowledge of the reference standard in all studies apart from one (99%) (Akarsu 2006). This occurred because the reference standard relied on the interpretation of the visual assessment to decide whether to complete an excavation of the site. The nature of the visual examination which relies on the examiner applying a pre-determined scale ensured that the threshold was pre-specified in all studies apart from one study which did not report any details of the scale used by examiners (Sidi 1988). Two studies that applied multiple index tests used independent examiners for each test. This occurred where two different visual scales were applied by different examiners (Braga 2010) and where other tests were image based and could be saved on a computer and assessed remotely at a different time point (Laitala 2017). 19 were unclear on their use of examiners and 46 failed to use independent assessors for each index test and potentially introduced bias. This did not result in a high risk of bias as the visual assessment was always completed prior to the other index tests and the results would therefore not have been influenced. 46 (70%) studies were judged to be of high concern regarding the index test. These were studies where the index test was carried out on extracted teeth rather than teeth in situ. As for the patient selection domain, the applicability concern for visual classification systems in particular relates to the fact that when applied to extracted teeth that have been cleaned and tried

and may be optimally placed for assessment that this context is not reflective of a typical clinical setting.

### Reference standard

Of the 67 included studies, 54 (81%) were at low risk of bias for the reference standard, of these, 10 studies were conducted in a clinical setting on teeth due for extraction, and received an histological reference standard (Braga 2010; Castilho 2016; Cinar 2013; Diniz 2012; Freitas 2016; Goel 2009; Kockanat 2017; Kucukyilmaz 2015; Rocha 2003; Teo 2014). Nine were at high risk of bias (Akarsu 2006; Bahrololoomi 2015; Bussanelli 2015; Huth 2010; Kim 2017; Laitala 2017; Novaes 2009; Novaes 2010; Sidi 1988) because of our concern that the reference standard may not correctly classify the enamel caries. This is due to the use of excavation of teeth identified as being severe caries which warrant restoration (Akarsu 2006; Bahrololoomi 2015; Huth 2010), using radiographs as a reference standard which are unlikely to correctly classify the target condition (Kim 2017; Laitala 2017; Sidi 1988) or using a visual examination after separation of the approximal surfaces (Bussanelli 2015; Novaes 2009; Novaes 2010). We were confident that 58 studies (87%) used a reference standard which would correctly identify enamel caries. It was unclear whether independent examiners assessed the reference standard in 36 studies however, we were satisfied that the reference standard examiner was independent in 27 studies. Concerns for applicability regarding the reference standard were low.

### Flow and timing

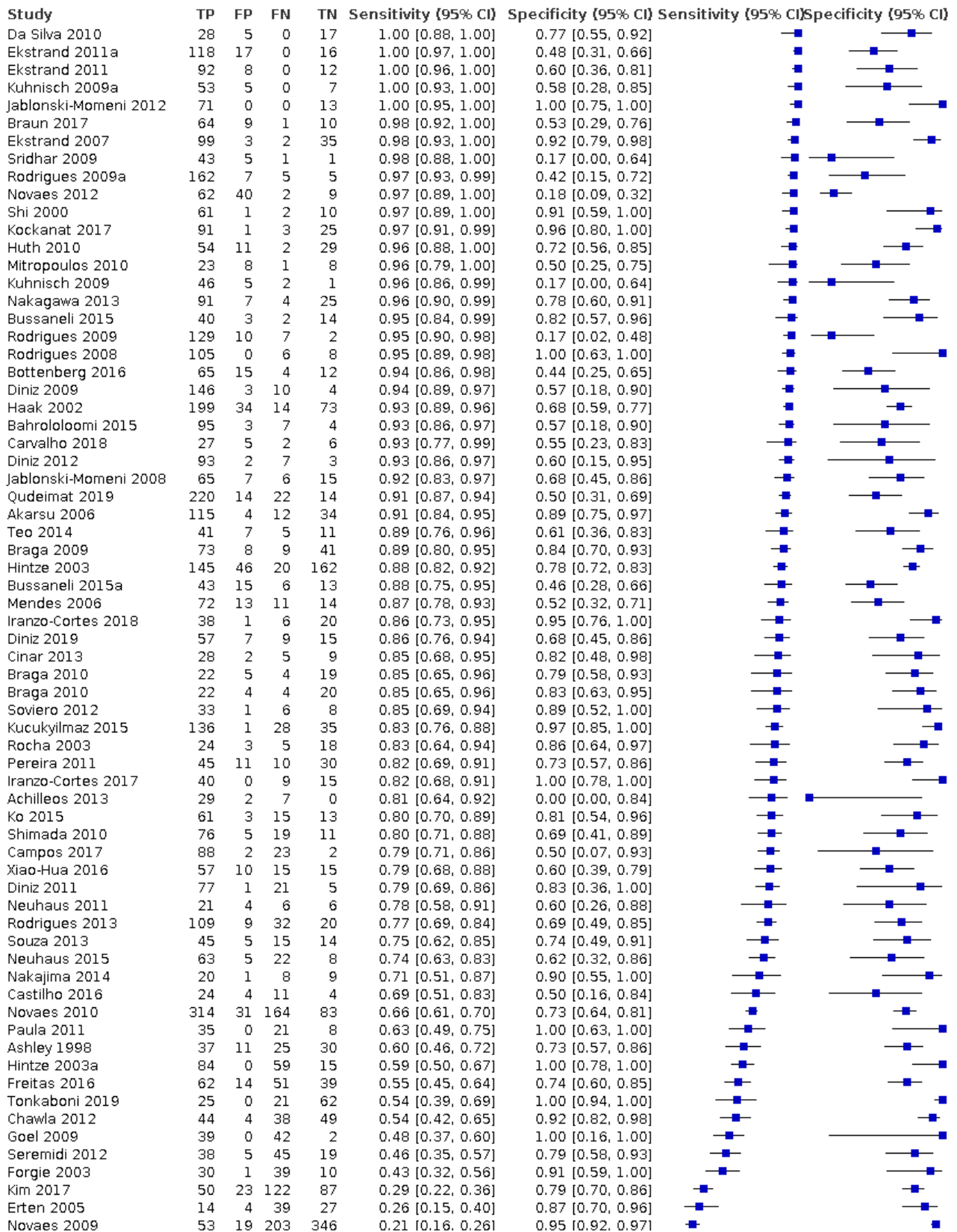
Fifty-nine studies were judged to be at low risk of bias for patient flow (88%), one study was judged to be unclear regarding the time interval between index test and reference standard as the teeth were monitored for natural exfoliation and histology performed after exfoliation (Freitas 2016). There was variation in the reference standards in three studies (Akarsu 2006; Bahrololoomi 2015; Huth 2010). Four studies failed to apply the reference standard to all included teeth (Jablonski-Momeni 2008; Kim 2017; Kockanat 2017; Shi 2000). Where studies clearly reported that teeth did not receive the reference standard as a result of breakages during the

sectioning procedure, then this did not affect the overall bias result for this domain.

### Findings

The point of assessment was the tooth surface, and some studies assessed multiple sites on the same surface. Visual assessment was used to detect early/enamel caries in 19,590 teeth or tooth sites with 35% prevalence enamel caries. 52 datasets reported one surface per tooth and 19 reported multiple sites on each tooth, of these 12 datasets investigated proximal surfaces (Braga 2009; Ekstrand 2011; Ekstrand 2011a; Freitas 2016; Haak 2002; Hintze 2003; Laitala 2017; Mitropoulos 2010; Novaes 2009; Novaes 2010; Sidi 1988; Soviero 2012). The principal findings of this review are reported for all included datasets assessing a visual classification system for detecting enamel caries, with no restriction on tooth surfaces, dentition, reference standard or prevalence of disease (Additional Table 4). The bivariate method was used for meta-analysis, the individual studies had sensitivities which ranged from 0.16 to 1.00 and specificities from 0 to 1.00. The coupled forest plots of sensitivity and specificity for each of the included datasets is illustrated in Figure 5. The overall summary sensitivity and specificity values were 0.86 (95% confidence interval (CI) 0.80 to 0.90) and 0.77 (95% CI 0.72 to 0.82) respectively. The diagnostic odds ratio (DOR) was 20.38 (95% CI 14.33 to 28.98). Figure 6 presents the summary receiver operating characteristic (SROC) plot with the summary point plotted along with the 95% confidence region and prediction region. The relatively narrow confidence intervals and confidence region for the summary sensitivity and specificity estimate are reflective of the volume of data in the analysis. The prediction region shows the range of likely values for a future individual study, and was very broad, covering the top half of the plot, and indicating substantial heterogeneity across the studies. Summary of findings 1 reports the results of the 67 included studies which generated 71 datasets. We rated the certainty of the evidence as low, and downgraded two levels in total, for risk of bias due to limitations in the design and conduct of the included studies, indirectness arising from the in vitro studies, and inconsistency of the results.

**Figure 5. Forest plot of all available datasets, sorted according to sensitivity.**

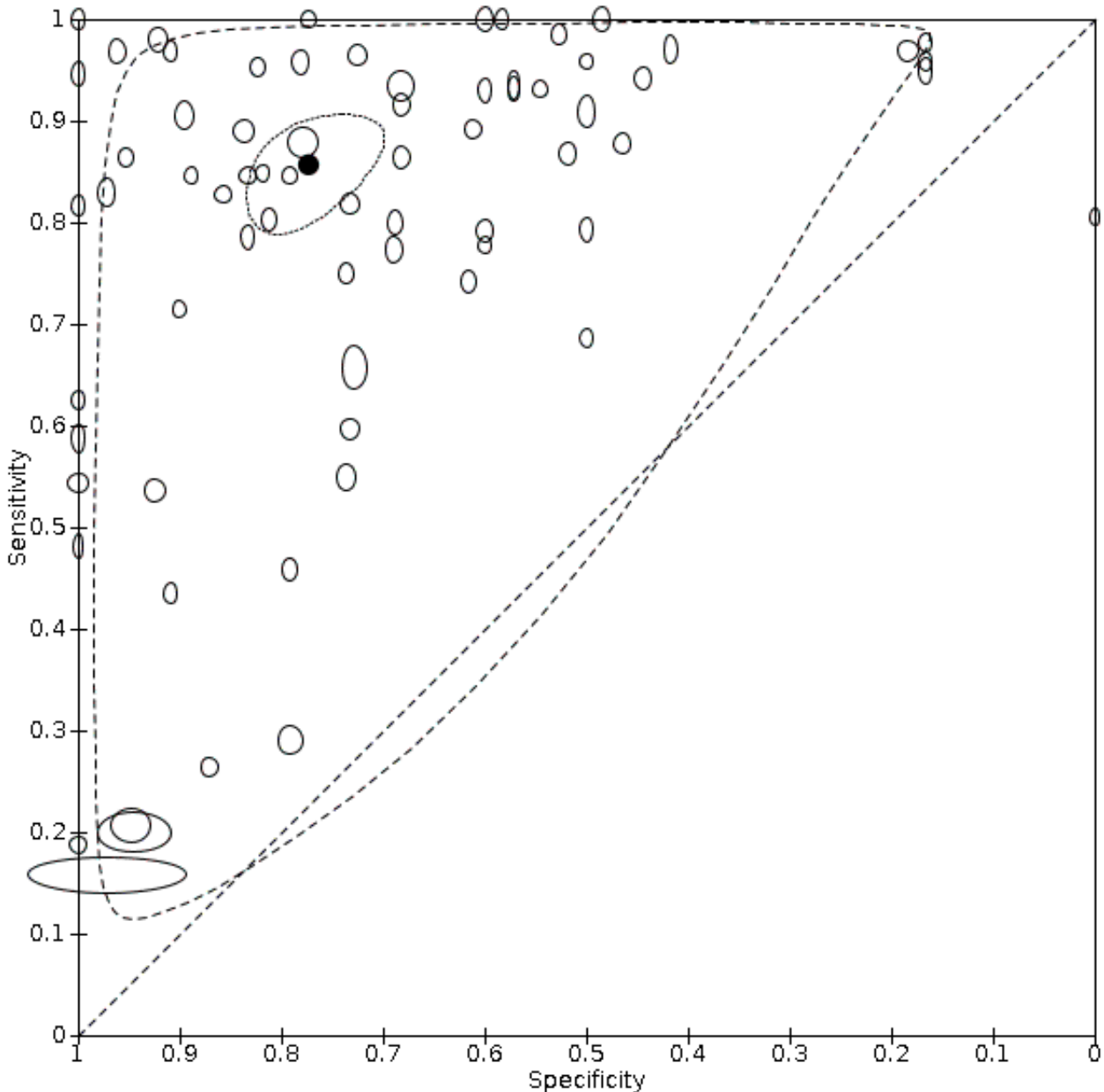




**Figure 5. (Continued)**

Nili 2017	30	23	122	87	0.29 [0.22, 0.36]	0.79 [0.70, 0.88]	
Erten 2005	14	4	39	27	0.26 [0.15, 0.40]	0.87 [0.70, 0.96]	
Novaes 2009	53	19	203	346	0.21 [0.16, 0.26]	0.95 [0.92, 0.97]	
Laitala 2017	83	93	332	1595	0.20 [0.16, 0.24]	0.94 [0.93, 0.96]	
Costa 2002	6	0	26	17	0.19 [0.07, 0.36]	1.00 [0.80, 1.00]	
Sidi 1988	45	255	240	8558	0.16 [0.12, 0.21]	0.97 [0.97, 0.97]	

**Figure 6. Summary sensitivity and specificity points of all included datasets with 95% confidence and prediction regions.**



We explored the comparative accuracy of the different classification systems (Figure 7). ICDAS (38 datasets, 6817 tooth surfaces) and ERK (15 datasets, 1306 tooth surfaces) were the

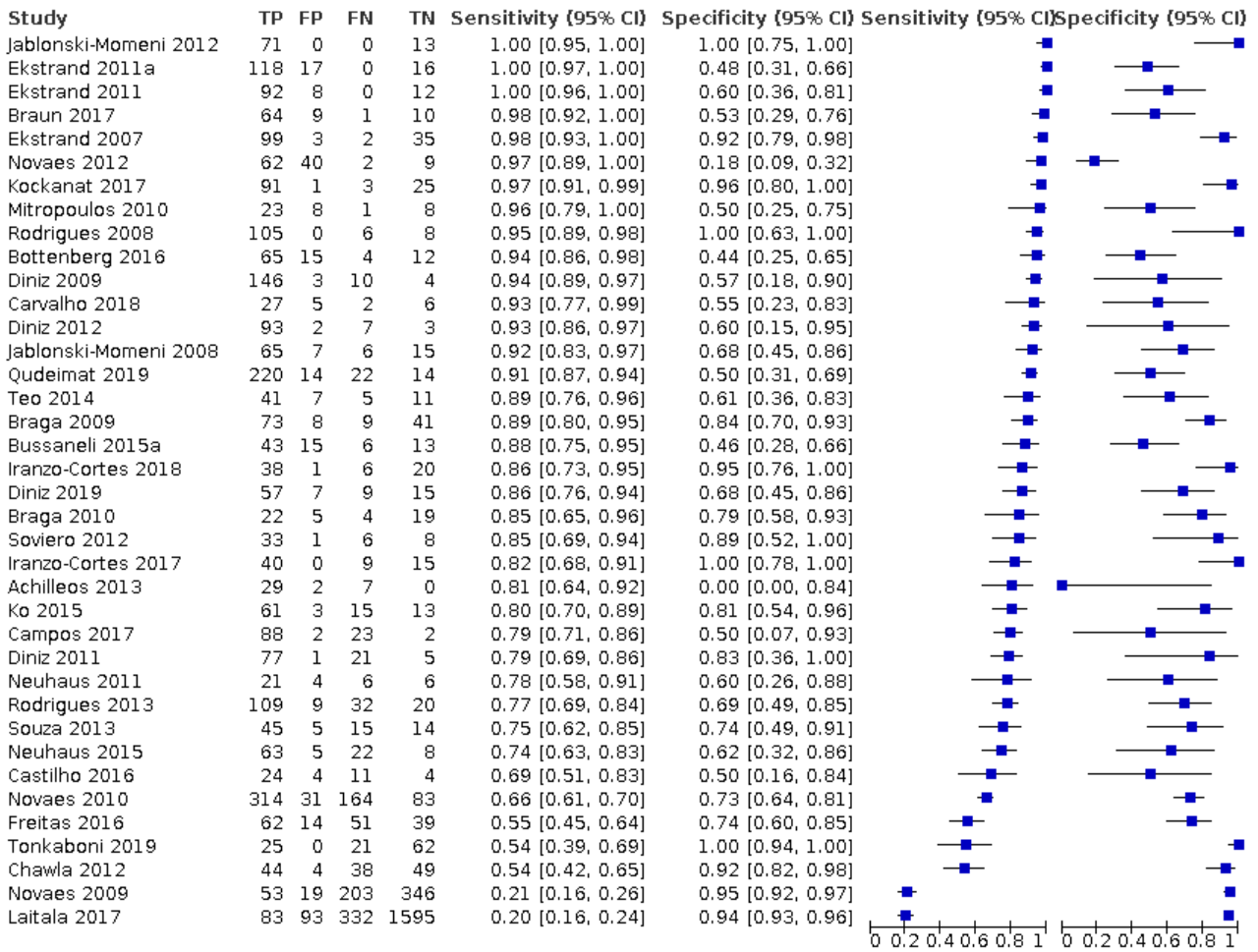
most commonly used classification systems. For the purposes of analysis we created a category of 'other visual' that comprised the remaining 18 datasets (11,467 surfaces) that either did not report

the classification system used, or reported the use of a unique classification system (e.g. Downer, Nyvad). We explored the effect of different classification systems by including a covariate for test type in the bivariate model to assess the effects on sensitivity or specificity or both. The summary points for sensitivity and specificity were 0.88 (95% CI 0.81 to 0.92) and 0.76 (95% CI 0.67 to 0.83) for ICDAS, 0.85 (95% CI 0.74 to 0.92) and 0.72 (95% CI 0.61 to 0.81) for ERK, and 0.81 (95% CI 0.65 to 0.91) and 0.84 (95%

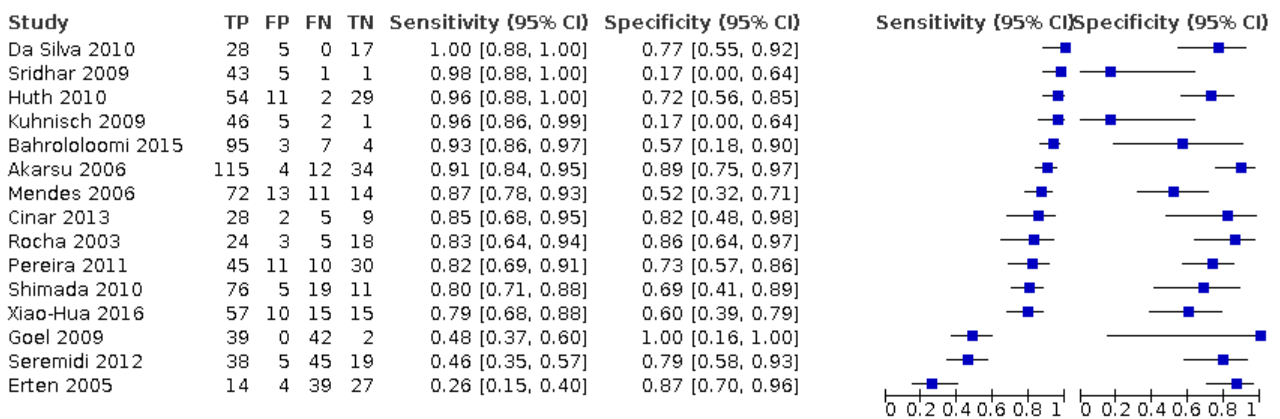
CI 0.73 to 0.91) for other visual. The addition of test type to the model did not result in any meaningful difference to the sensitivity or specificity estimates ( $\text{Chi}^2(4) = 3.78$ ,  $P = 0.44$ ) (Figure 8). The prediction regions for each of the different classification systems were very broad, covering the top half of the plot, indicating substantial heterogeneity across the studies within each of the classification systems.

**Figure 7. Forest plot grouped according to visual classification system (ICDAS, ERK, and other visual examinations) sorted according to sensitivity.**

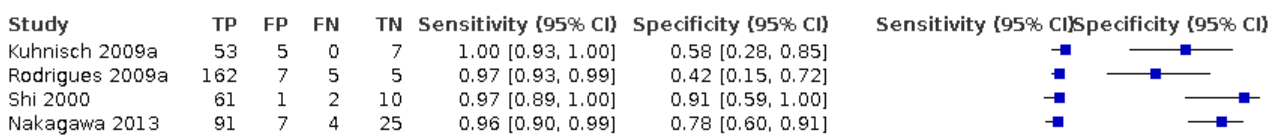
ICDAS



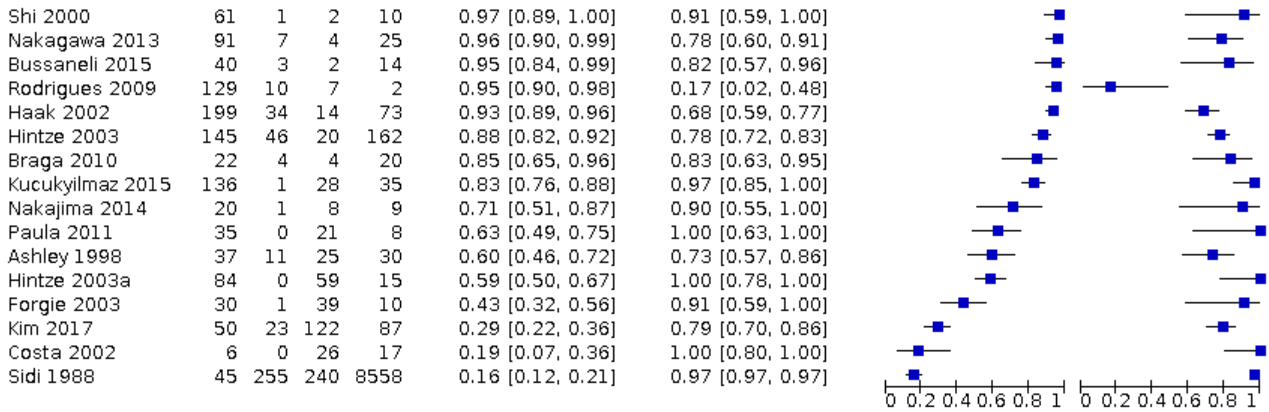
ERK



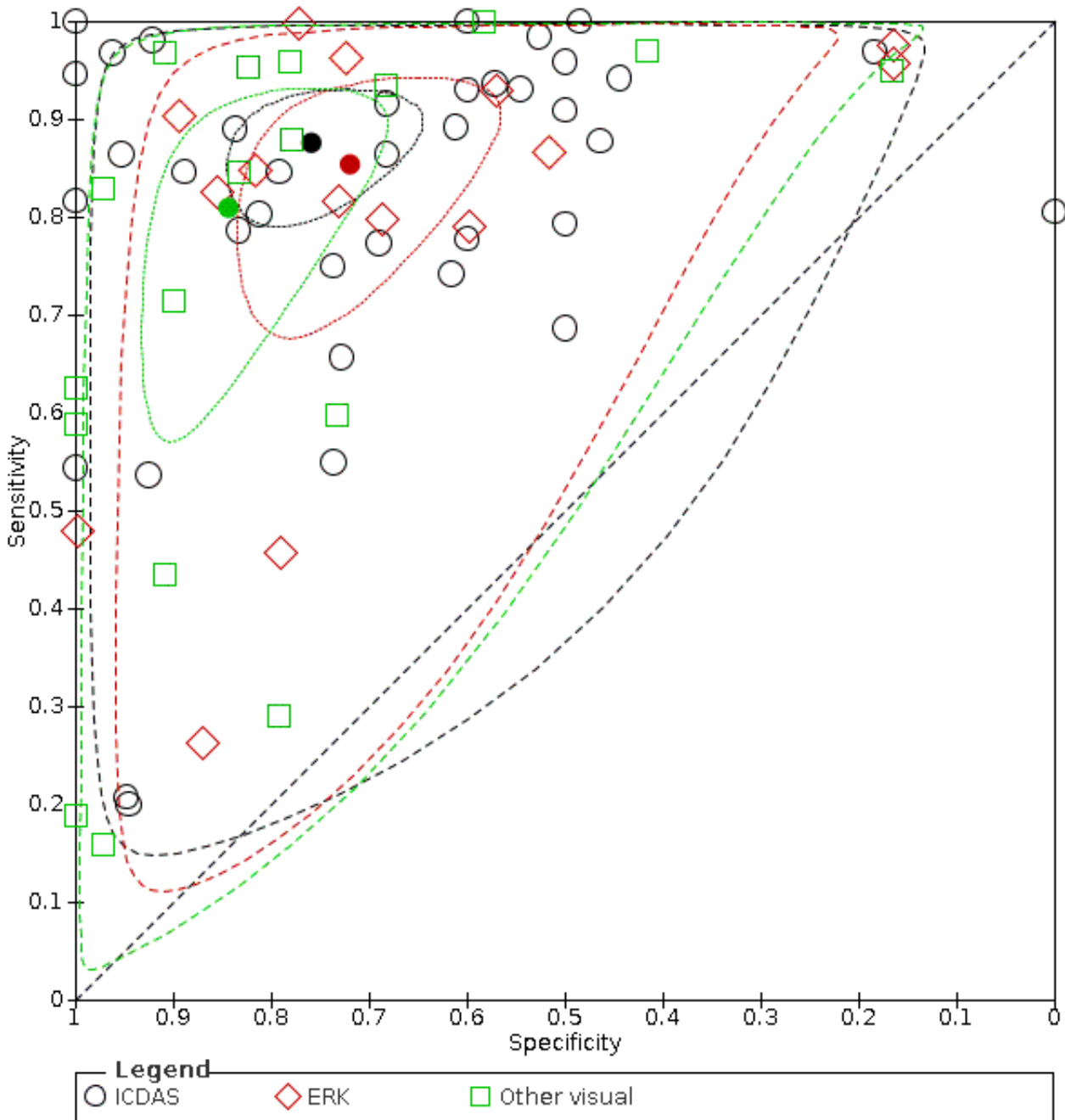
Other visual



**Figure 7. (Continued)**



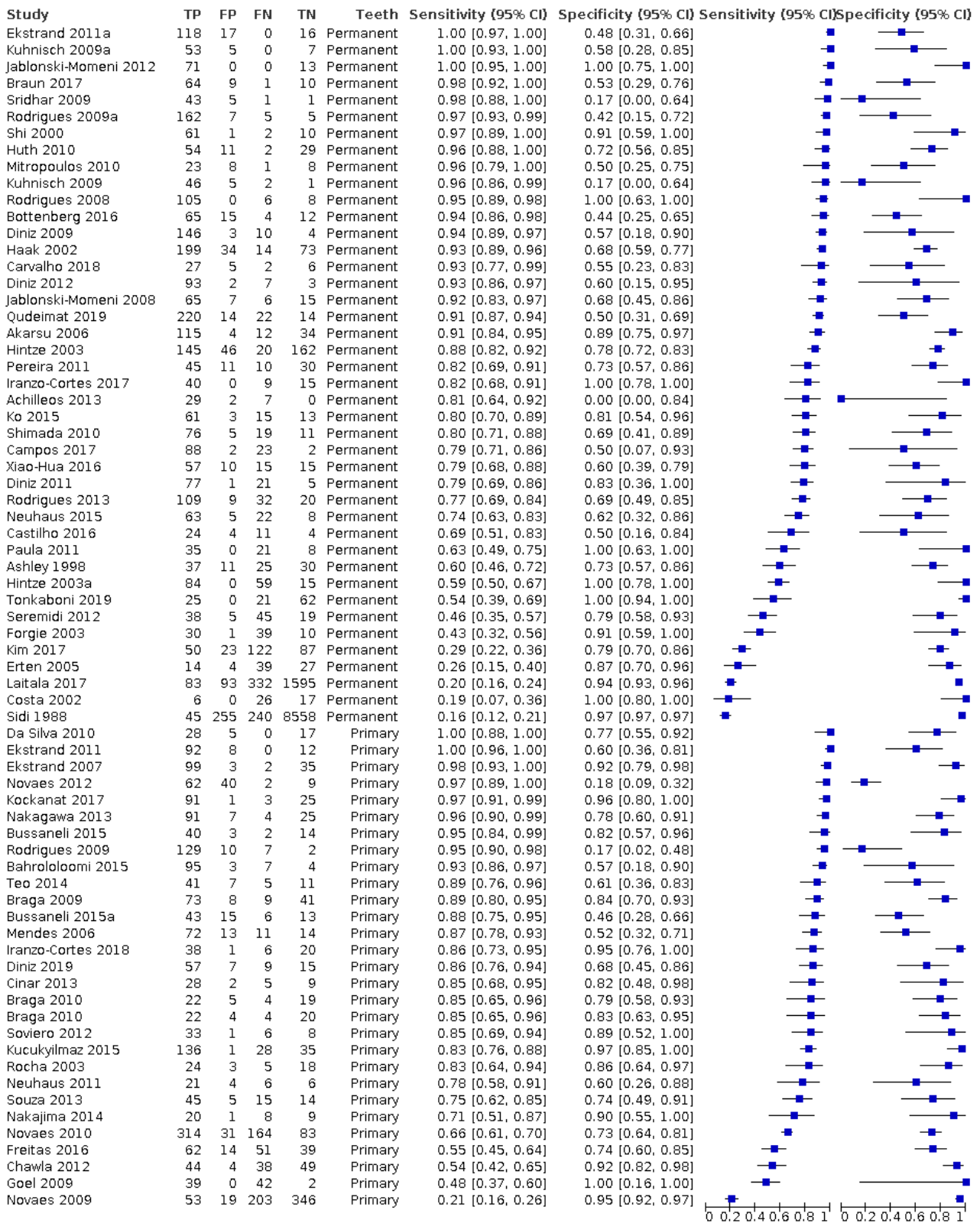
**Figure 8. Summary sensitivity and specificity points by ICDAS, ERK, and other visual classification system with 95% confidence and prediction regions.**



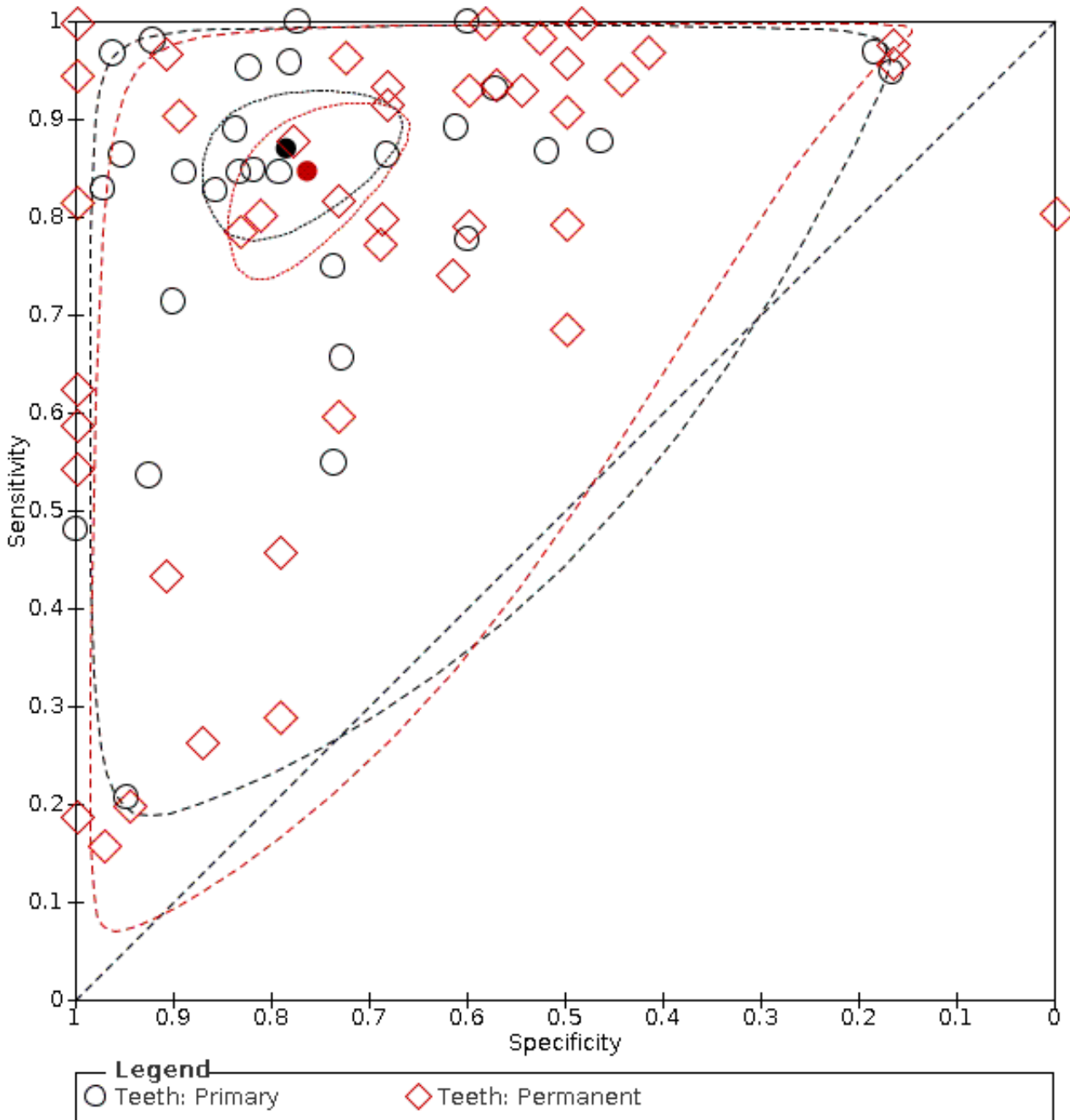
A coupled forest plot (Figure 9) illustrates the sensitivity and specificity of datasets reporting on the primary (29 datasets, 3705 tooth surfaces) and permanent dentition (42 datasets, 15,885 tooth surfaces) and sorted by sensitivity. Summary points for each dentition are plotted in Figure 10. The summary estimates from the bivariate model are sensitivity 0.87 (95% CI 0.80 to 0.92) and specificity 0.79 (95% CI 0.70 to 0.85) for the primary dentition, and

sensitivity 0.85 (95% CI 0.77 to 0.90) and specificity 0.76 (95% CI 0.69 to 0.83) for the permanent dentition. The addition of dentition to the model did not result in any meaningful difference to the sensitivity or specificity estimates ( $\text{Chi}^2(2) = 0.90, P = 0.64$ ). The prediction regions for each dentition were very broad, covering the top half of the plot, indicating substantial heterogeneity across the studies within each dentition.

**Figure 9. Forest plot grouped according to primary or permanent dentition, sorted according to sensitivity.**



**Figure 10. Summary sensitivity and specificity points by primary and permanent dentition with 95% confidence and prediction regions.**



Results according to visual classification system and dentition are presented in [Summary of findings 2](#).

**Investigations of heterogeneity**

Tests for heterogeneity were applied across all available studies as there was no evidence of a difference between the classification systems used. Meta-regression was used to explore the potential sources of heterogeneity.

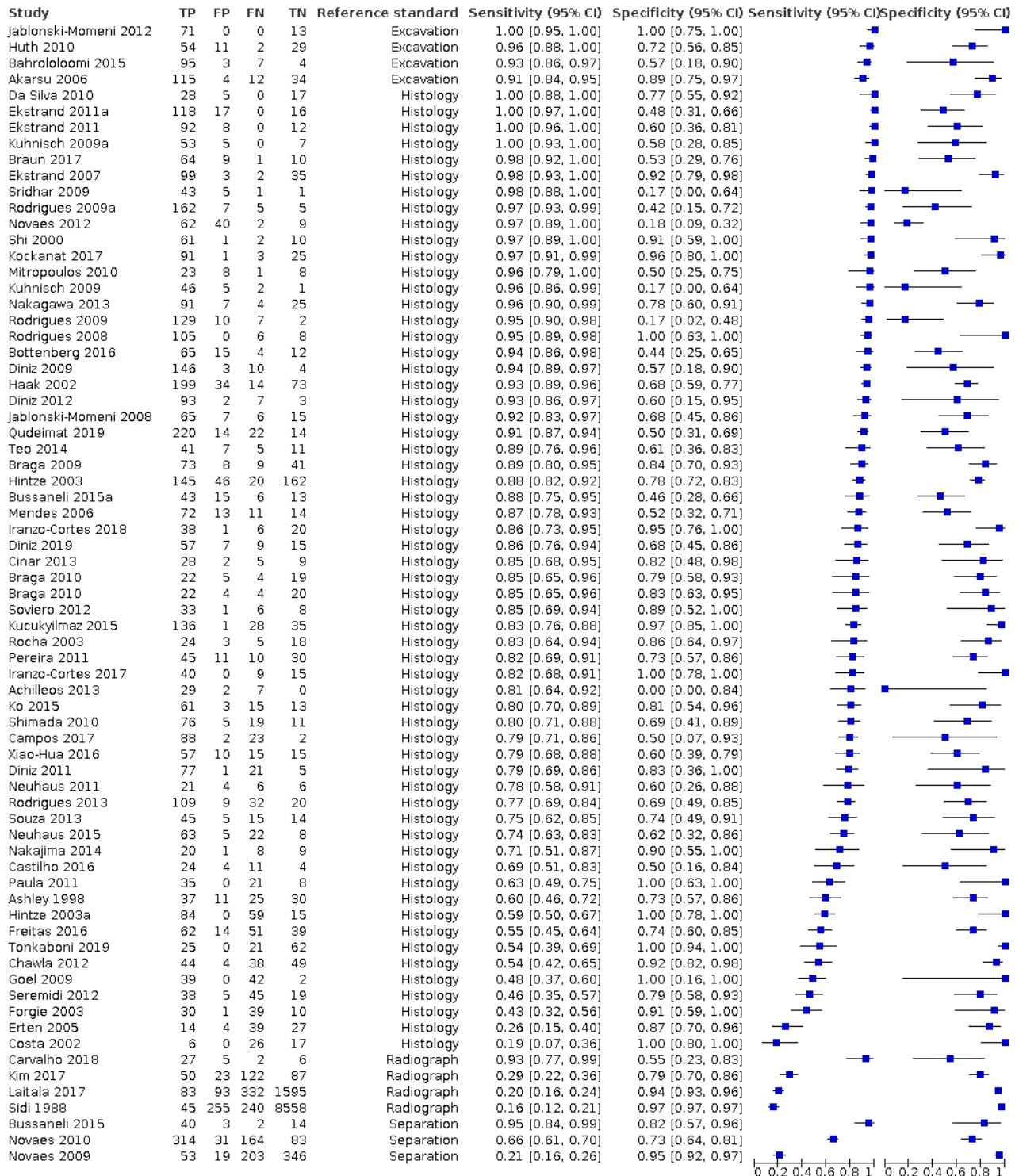
**Reference standard**

Sixty datasets (85%) reported the use of a histological reference standard, 49 on previously extracted teeth (in vitro studies), and 11 on teeth indicated for extraction or exfoliation (in vivo studies). Four datasets used either bitewing radiographs or microcomputed tomography (microCT) on previously extracted teeth, three datasets separated teeth in situ to improve the clinicians view of the approximal surface, and four studies chose to excavate the teeth to determine the severity of caries, one of these

was performed on previously extracted teeth (Figure 11; Figure 12). Due to the paucity of studies in the excavation, radiograph, and separation reference standard categories, these categories were combined into a non-histology category for the purpose of meta-regression. Summary sensitivity and specificity estimates from

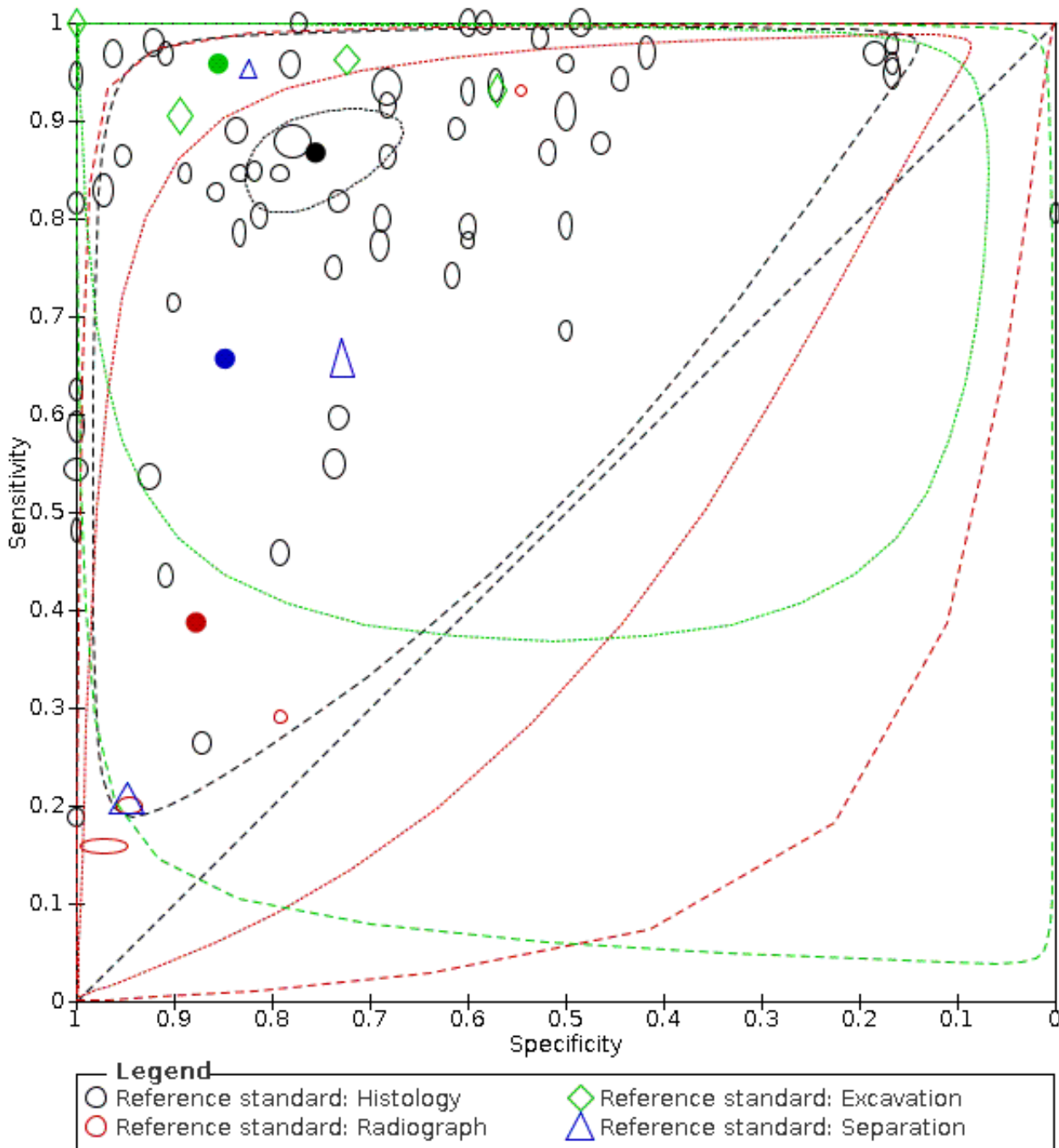
studies using a histological or non-histological reference standard were similar: histology sensitivity 0.87 (95% CI 0.82 to 0.90) and specificity 0.76 (95% CI 0.69 to 0.81); non-histology sensitivity 0.78 (95% CI 0.48 to 0.93) and specificity 0.86 (95% CI 0.76 to 0.92);  $\chi^2(2) = 3.28, P = 0.19$ .

Figure 11. Forest plot grouped by reference standard, sorted according to sensitivity.





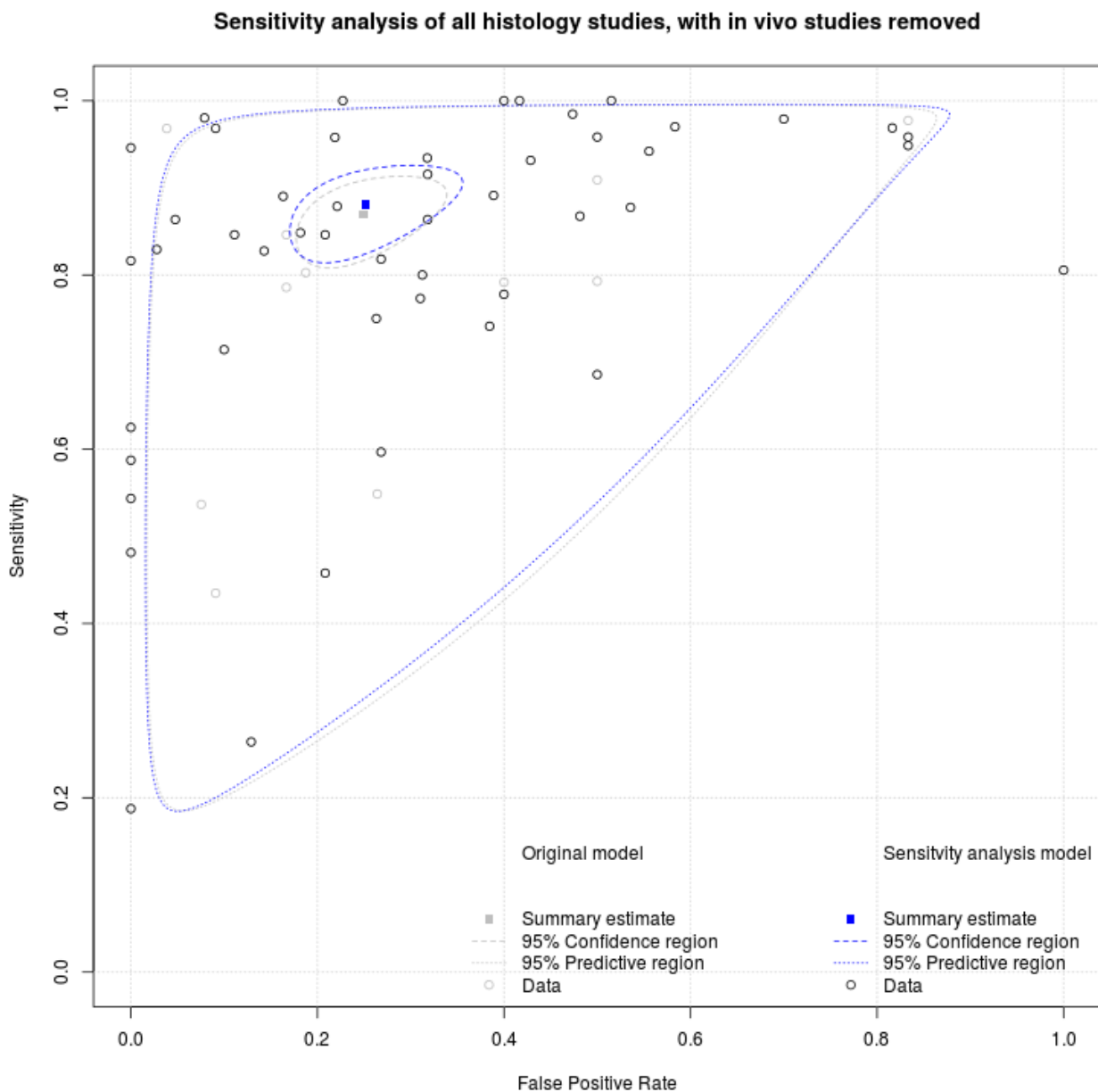
**Figure 12. Summary sensitivity and specificity points by reference standard with 95% confidence and prediction regions.**



It was stated in the protocol that in vitro and in vivo studies would be compared, since 49 of the 60 studies using histology were in vitro this analysis has not been completed. A comparison of the histology studies comparing the in vitro and in vivo studies has been performed as a sensitivity analysis to assess whether the

additional in vivo studies make a meaningful difference to this group of studies. Figure 13 confirms a minimal change to the estimate of summary sensitivity and specificity when the in vivo studies are removed.

**Figure 13. Sensitivity analysis of studies using a reference standard only, with in vivo studies removed.**

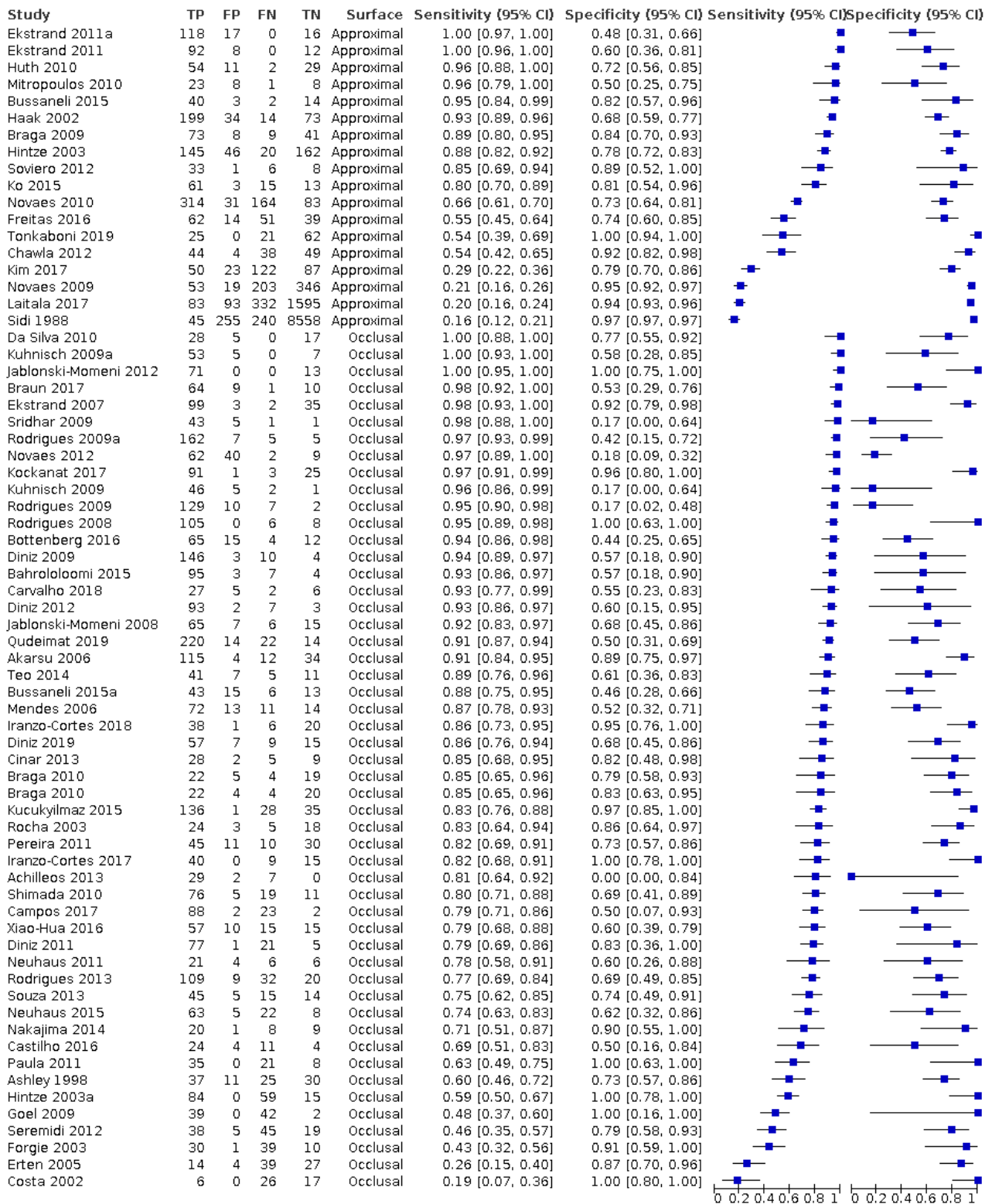


**Surface**

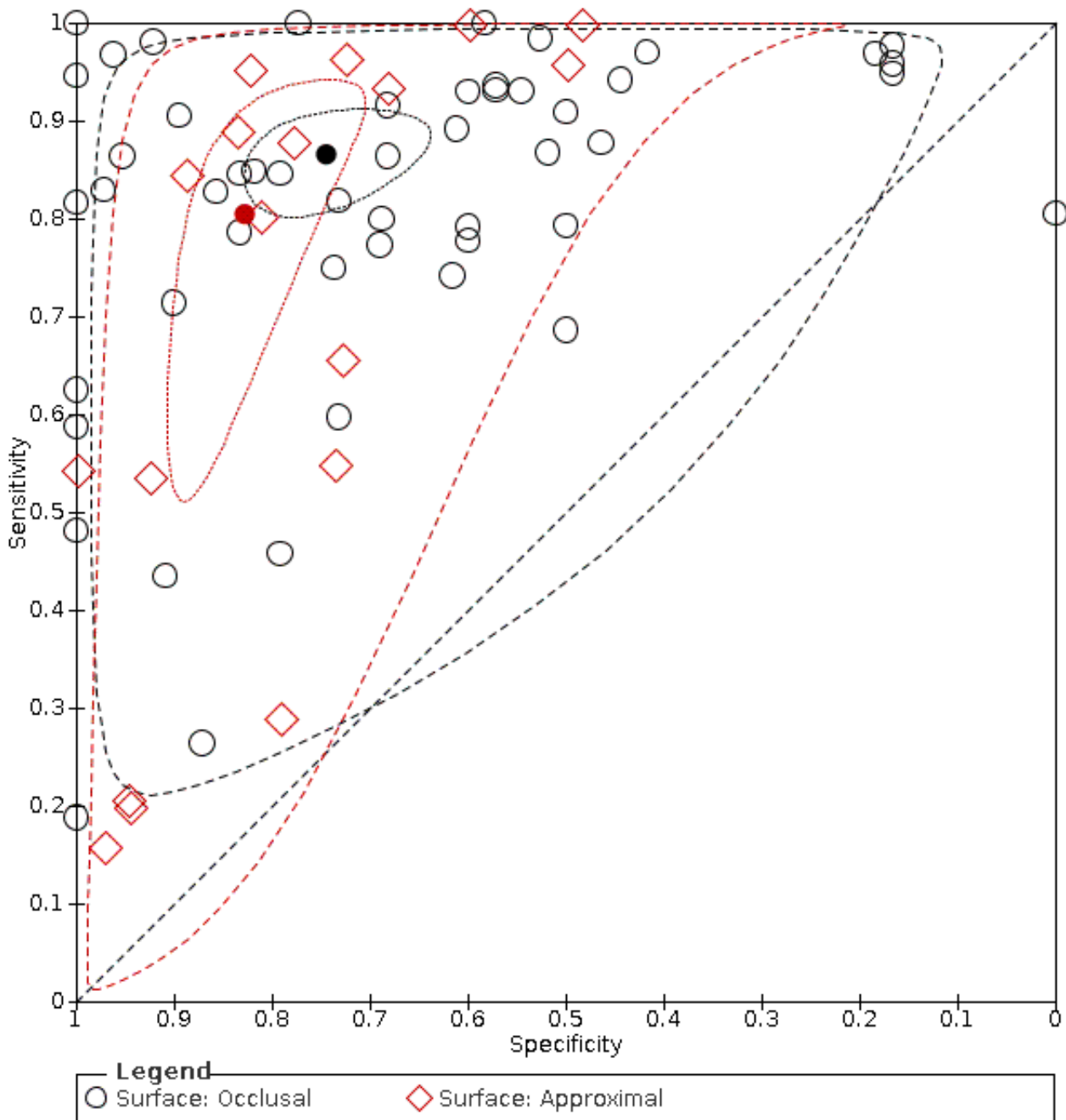
Included studies generated 51 occlusal (72%) and 18 approximal datasets, the remaining two datasets reported on smooth surfaces (Figure 14). Of the 18 approximal surface investigations eight datasets were completed in a clinical setting (Bussaneli 2015; Freitas 2016; Huth 2010; Kim 2017; Laitala 2017; Novaes 2009; Novaes 2010; Sidi 1988), four attempted to recreate the abutment of teeth in the oral cavity (Braga 2009; Huth 2010; Ko 2015; Tonkaboni 2019), and six assessed the surface without setting teeth in a model which would recreate this (Ekstrand 2011; Ekstrand 2011a; Haak 2002; Hintze 2003; Mitropoulos 2010; Soviero 2012). There

is a heterogeneity in sensitivity and specificity for both groups. Due to the sparsity of studies evaluating the smooth surfaces this meta-analysis was restricted to the 69 datasets reporting values for the occlusal and approximal surfaces only. The bivariate meta-analysis produced summary estimates of sensitivity and specificity for occlusal surfaces: sensitivity 0.87 (95% CI 0.82 to 0.90) and specificity 0.74 (95% CI 0.66 to 0.81); and approximal surfaces: sensitivity 0.80 (95% CI 0.60 to 0.92) and specificity 0.83 (95% CI 0.74 to 0.89) (Figure 15). The addition of tooth surface to the model did not result in any meaningful difference to the sensitivity or specificity estimates (Chi<sup>2</sup> (2)= 3.09, P = 0.21).

Figure 14. Forest plot grouped by tooth surface, sorted according to sensitivity.



**Figure 15. Summary sensitivity and specificity points by tooth surface with 95% confidence and prediction regions.**

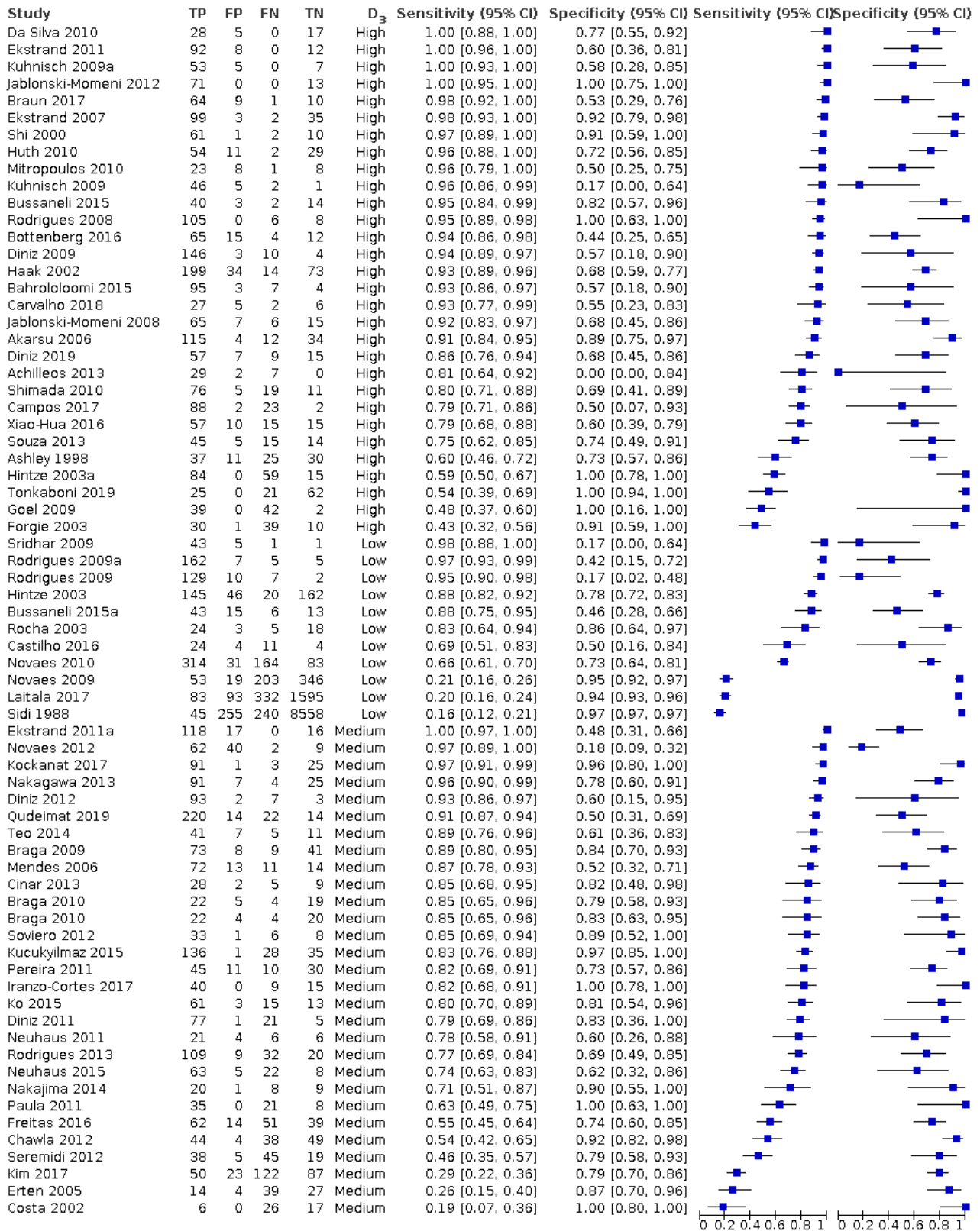


**Prevalence of D<sub>3</sub> disease**

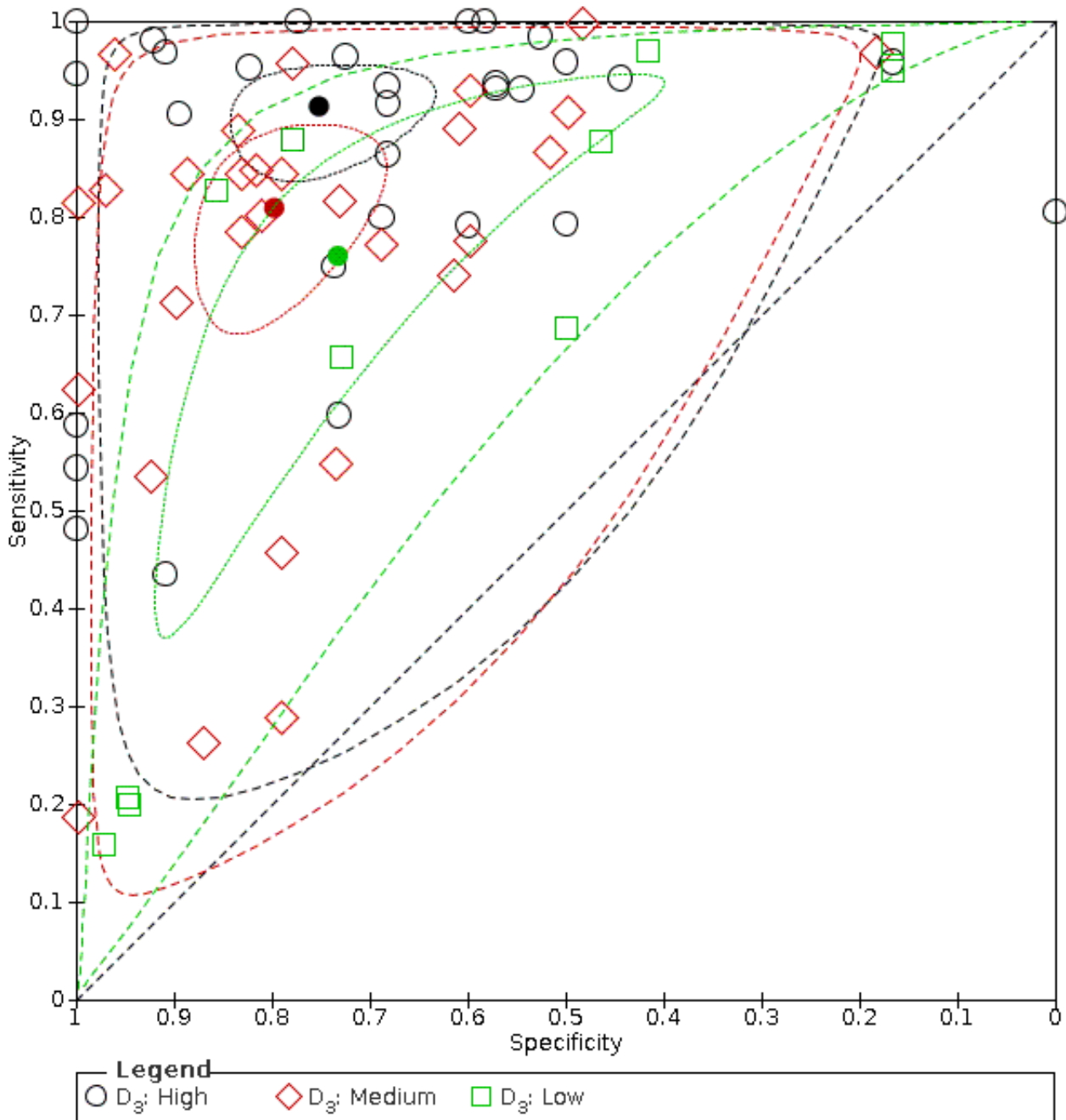
The datasets were categorised according to the prevalence of dentine caries with low prevalence (0% to 14%), medium (15% to 34%) and high ( $\geq 35\%$ ), as per the other reviews in this series (Figure 16). Where studies did not report the prevalence of dentine caries an estimation of the level was made from the description in the paper and the reporting of the enamel prevalence. We analysed 31 (44%) datasets that reported a high prevalence, 28 (40%) datasets

that reported a medium prevalence, and 11 (16%) datasets that reported a low prevalence. The prevalence of dentine caries was not reported and could not be estimated in one dataset. As may be expected, estimates of sensitivity are most consistent in the high prevalence group and range from 0.43 to 1.00. The summary points are illustrated in Figure 17. From the bivariate model with a covariate for prevalence category and assuming common variances the estimates for sensitivity and specificity are.

**Figure 16. Forest plot grouped by prevalence of dentine caries (low: 0% to 14%, medium: 15% to 34%, and high: ≥ 35% prevalence) sorted according to sensitivity.**



**Figure 17. Summary sensitivity and specificity points by prevalence of dentine caries (low: 0% to 14%, medium: 15% to 34%, and high: ≥ 35% prevalence) with 95% confidence and prediction regions.**



- Low prevalence: 0.76 (95% CI 0.50 to 0.91) and 0.73 (95% CI 0.51 to 0.88).
- Medium prevalence: 0.81 (95% CI 0.72 to 0.88) and 0.80 (95% CI 0.71 to 0.86).
- High prevalence: 0.91 (95% CI 0.86 to 0.95) and 0.75 (95% CI 0.66 to 0.82).

sensitivity increases with prevalence of caries into dentine ( $\text{Chi}^2(10) = 20.34, P = 0.03$ ). Prevalence values are likely to encompass many different underlying factors, and so this result should be interpreted with caution.

In conclusion, only one of the potential sources of heterogeneity specified a priori appeared to explain the observed heterogeneity.

The addition of  $D_3$  prevalence category to the model indicated differences in sensitivity or specificity or both, suggesting that

## Sensitivity analysis

Sensitivity analysis was proposed restricting studies to:

- low risk of bias on the inclusion criteria for caries threshold;
- low risk of bias for the index test domain;
- low risk of bias for the reference standard domain; and
- low prevalence of dentine caries (i.e. < 35%).

The poor overall quality of many of the studies precluded any meaningful sensitivity analysis, as to do so would have meant discarding a substantial number of included studies from the analysis. For example, many studies failed to clearly report the level of cavitation that was present in the eligible population from which participants were recruited. Descriptions of the inclusion criteria used in many of the studies were often poorly reported.

The overall estimates for all 71 datasets were 0.86 (95% CI 0.80 to 0.90) for sensitivity and 0.77 (95% CI 0.72 to 0.82) for specificity.

Ten datasets (Bahrololoomi 2015; Braga 2009; Da Silva 2010; Goel 2009; Mendes 2006; Neuhaus 2011; Novaes 2012; Rocha 2003; Shi 2000; Sidi 1988) were judged as high or unclear risk of bias for the index test. Removing these datasets from the analysis resulted in estimates of sensitivity and specificity of 0.86 (95% CI 0.81 to 0.90) and 0.76 (95% CI 0.70 to 0.81), a marginal change to the overall estimates.

Thirteen datasets (Akarsu 2006; Bahrololoomi 2015; Bottenberg 2016; Bussaneli 2015; Bussaneli 2015a; Huth 2010; Kim 2017; Laitala 2017; Neuhaus 2011; Novaes 2009; Novaes 2010; Sidi 1988; Tonkaboni 2019) were judged as high or unclear risk of bias for the index test. Removing these datasets from the analysis resulted in estimates of sensitivity and specificity of 0.88 (95% CI 0.83 to 0.92) and 0.75 (95% CI 0.67 to 0.81), a marginal change to the overall estimates.

The reported or imputed prevalence of dentine caries was low (< 15%) in 11 studies (Bussaneli 2015a; Castilho 2016; Hintze 2003; Laitala 2017; Novaes 2009; Novaes 2010; Rocha 2003; Rodrigues 2009; Rodrigues 2009a; Sidi 1988; Sridhar 2009). The sensitivity and specificity estimates of these datasets was 0.76 (95% CI 0.51 to 0.91) and 0.73 (95% CI 0.51 to 0.88), lower than the overall estimates.

## DISCUSSION

### Summary of main results

The main findings of the review are as follows.

- **For all included studies, the summary estimates were sensitivity of 0.86 (95% confidence interval (CI) 0.80 to 0.90) and specificity of 0.77 (95% CI 0.72 to 0.82), low-certainty evidence.** In a hypothetical cohort of 1000 tooth sites or surfaces with an illustrative prevalence of enamel caries of 28% (Steele 2011), these results would mean 40 tooth sites or surfaces not being identified as having early caries when caries was present (false negatives) and 163 tooth sites or surfaces being identified as having caries when they did not (false positives) (Summary of findings 1).
- **The effect of test type and dentition on summary estimates was minimal ( $\text{Chi}^2(4) = 3.78, P = 0.44; \text{Chi}^2(2) = 0.90, P = 0.64$ ).** The centre of the observed data was similar for the different classification systems and for the primary and

permanent dentition (Summary of findings 2). The prediction regions indicated substantial heterogeneity across the studies within each of the classification systems and dentitions.

- **Only one of the investigated potential sources of heterogeneity was able to explain the observed heterogeneity.** Of all the sources investigated only prevalence of caries into dentine provided results beyond those expected by chance ( $\text{Chi}^2(4) = 9.64, P = 0.04$ ), and suggested that sensitivity increases with prevalence. Intuitively this is likely as where advanced lesions are present in the sample they will be easier to detect.

In terms of misclassification following the use of these visual classification systems, the consequences can be considered to be quite low impact given the clinical considerations, although it should be acknowledged that financial and resource costs will be incurred. Where early caries is missed, this is most likely be early in the disease continuum and given the slow progressing nature of caries, particularly when limited to the enamel of the tooth, and the regular recall interval of many dental patients, we could be satisfied that the lesion would be identified at the next appointment. False-positive classifications could potentially result in care where it was not required and would probably be limited to the application of fluoride varnish. In such a scenario there are cost implications, but there is low potential for an adverse event for the patient. It is important that results are interpreted with caution given the considerable unexplained heterogeneity that is reflected in the 95% prediction region. So whilst on average the confidence intervals indicate a reasonable performance, they do not reflect the confidence that one can have in the accuracy of visual assessments using these classification systems as there is considerable unexplained heterogeneity evident, and the prediction regions in which the sensitivity and specificity of a future study should lie are very broad. This is perhaps not surprising, as whilst visual classification systems may have multiple categories for describing the severity of carious lesions, these are still subject to individual interpretation and recall over time.

There is some similarity in the classification of early lesions across the different systems. For example, the International Caries Detection and Assessment System (ICDAS) codes 0, 1, and 2 are identical to those in the Ekstrand-Ricketts-Kidd (ERK) system, and it is only at more severe levels of caries that these two systems differ in approach, where ERK code 3 is divided between codes 3 and 4 in ICDAS, and cavitated lesions (ERK 4) are divided between 5 and 6 in ICDAS depending on the size of the cavity. The ICDAS and ERK studies have been conducted mainly by dentists that have been involved in the development of the classification systems, or dentists in secondary care or academic institutes that have had significant training and calibration in the use of the systems and the transferability into general practice has yet to be investigated. The time taken for such meticulous examination of clean dry teeth that these systems require is a consideration for many dentists who work under remuneration systems that inadequately reward the additional time these take.

Most included studies were of an in vitro design using previously extracted teeth and a reference standard of histology on occlusal surfaces. The investigation of approximal surfaces is somewhat flawed because when detecting approximal caries, the area beneath the contact of two teeth is not usually able to be seen directly. In this situation dentists look for two things: the lesion

at the surface of the tooth if it extends sufficiently around the buccal and lingual curvatures of the proximal surface, or shadows through the marginal ridge. As such, these lesions are generally more extensive and false positives (low specificity) less likely. One concern for the approximal dataset is the use of extracted teeth without simulating the proximal contact points. This is not a fair representation of the true clinical scenario and this concern was taken into account in the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) judgements. The lower specificity for occlusal surfaces compared to approximal surfaces may be explained in part by staining which is commonly present in the pits and fissures of the occlusal surface. Differentiating staining from early brown spot (enamel) lesions is notoriously difficult for dentists (Cortes 2003).

The included studies allowed us to evaluate the diagnostic test accuracy of visual or visual-tactile classification systems for the detection of early or non-cavitated caries, with particular focus on early stage caries in the enamel of the tooth. The motivation for the choice of target condition was that by detecting caries at this early stage, preventive or minimally invasive interventions can be applied which may avoid the need for future restorations. The part of the diagnostic pathway being investigated in the included studies was the detection element, it was not possible to assess diagnosis as no studies reported the inclusion of additional information for diagnosis. Many studies were excluded because they did not report data at this enamel threshold or because they deliberately included more advanced lesions into dentine. Other studies met the criteria for inclusion but did not present data in a form that enabled us to include them in the meta-analysis.

Overall risk of bias was low across all domains apart from participant selection. Where non-random or non-consecutive selection of the study sample occurs, teeth could be selected where disease is easier or more difficult to identify and potential upward or downward bias introduced. The index tests and reference standards were generally well conducted. The primary reason for bias occurred when multiple index tests were performed (other than visual) and there may have been potential for these to influence the visual assessment. However, it was typical for the visual test to be completed first so this was unlikely. Reference standards apart from histology were recorded as representing high risk of bias as we could not be confident that the true diagnosis was made. This is problematic for caries studies of detection and diagnosis. Enhanced visual examination was accepted as a reference standard only in studies investigating approximal surfaces and where the reference standard was performed after tooth separation. In effect, these studies examined the potential for tooth separation to detect caries. The validity of the use of tooth separation is questionable as it is unlikely to be undertaken in clinical practice, and this was reflected in the risk of bias and applicability judgements. Our Cochrane diagnostic test accuracy (DTA) Review of radiographs published as part of this suite of reviews (Walsh 2021) has shown radiographs to be poor at the detection of enamel caries. Studies using radiographs as a reference standard were eligible for inclusion in this review in order to facilitate the inclusion of in vivo studies, however only five studies used radiographs as a reference standard. The final reference standard was excavation. This reference standard also resulted in a high risk of bias judgement, as only those teeth detected as having caries into dentine were allocated to receive the excavation assessment.

The large number of studies that were performed in vitro on previously extracted teeth resulted in judgements of high concern for applicability for the participant selection and index test domains. The availability of a suitable reference standard applicable to non-diseased tissue is a particular challenge for studies in this area. The only true way to confirm enamel caries is to section teeth and complete a histological assessment, however it is not practical to complete this on routinely attending patients in the clinical setting, as it would be unethical to extract healthy teeth or those with early caries lesions or easily restorable teeth. This challenge was addressed in some studies with the inclusion of participants who were identified as requiring non-caries related extractions, typically for orthodontic reasons, third molar extractions, or those participants whose primary teeth were scheduled to exfoliate. There are associated concerns regarding the generalisability of results from these studies: in the orthodontic scenario the immature permanent dentition may be only recently erupted and therefore limited in prevalence of caries; exfoliating primary teeth are not necessarily generalisable to the adult population; and third molar teeth are not necessarily representative of other molar teeth as they are difficult to clean and often have thick plaque deposits leading to widespread lesions which are easier to see.

### Strengths and weaknesses of the review

The strengths of this review are an extensive and comprehensive electronic search strategy with no limitations on publication status or language, a thorough application of the QUADAS-2 methodology carried out independently and in duplicate for all included studies, a robust and reproducible statistical analysis using the most current diagnostic test accuracy methods, and a grading of the evidence to guide clinicians and guideline developers.

Due to the nature of the visual examination there was always a consistent threshold applied by the studies. Although different scoring systems were applied it was always conceivable for examiners to categorise teeth or sites as sound or no caries versus any level of caries. We were therefore able to complete a bivariate meta-analysis which provided us with summary sensitivity and specificity results, with 95% confidence and prediction regions.

Our Cochrane Review builds on existing literature, notably the Bader 2002 review which included 39 studies without performing a meta-analysis and was not able to offer a firm conclusion due to the poor quality of available studies and high degree of heterogeneity in study design. It also only included histological reference standard. More recently Gimenez 2015 completed a review and meta-analysis of 102 included studies with a different disease threshold of advanced caries and a more limited synthesis of results. Despite these differences the conclusions are similar to our review where they state that "visual inspection presents good accuracy in the detection of carious lesions in primary and permanent teeth." Our review has allowed for a more in depth investigation of the evidence at the initial caries level, with additional caveats regarding interpretation on account of the observed heterogeneity.

The grouping of unknown scoring systems into a group labelled as 'other' could be considered a weakness of the review. Potentially these studies cover a spectrum of methods including complex classification systems and basic visual examinations more in line with aspects of the typical clinical examination routinely



performed by general dental practitioners. A comparison between a routine examination and more complex assessments would be beneficial to determine clinical implications for accuracy, health- and resource-outcomes. The application of non-histological reference standards should also be taken into consideration when interpreting the results of this review, although these are limited in number in this review.

### Applicability of findings to the review question

There are concerns that many of the included studies in this review are completed in an *in vitro* rather than a clinical setting and so may not be representative of the visual examination performed by a general dental practitioner in a clinical setting. Until a less imperfect reference standard appropriate for *in vivo* clinical studies is developed or more widely available, this is likely to be the status quo. The use of 3D technology such as microcomputed tomography (microCT) and cone beam computed tomography (CBCT) may go some way to improve upon these concerns in future studies.

## AUTHORS' CONCLUSIONS

### Implications for practice

The summary sensitivity and specificity points suggest that the visual classification systems evaluated in this review are reasonable at detecting early caries. However, it is important to remember that the confidence and prediction regions for the summary sensitivity and specificity points differ substantially. Whilst on average the confidence intervals indicate a reasonable performance, they should be interpreted in conjunction with the prediction regions that illustrate the unexplained heterogeneity; the prediction regions in which the sensitivity and specificity of a future study should lie are very broad. This is perhaps not surprising, as whilst visual classification systems may have multiple categories for describing the severity of carious lesions, these are still subject to individual interpretation and recall over time. Perhaps the greatest concern is that the methods investigated in this review involve detailed scoring systems which may be challenging to incorporate into the routine dental examinations conducted by general dental practitioners, although there is research available to suggest that this is achievable (Ormond 2010). The largest impact of this review could be the encouragement for general dental practitioners to adopt the rigorous methods of the International Caries Detection and Assessment System (ICDAS) or Ekstrand-Ricketts-Kidd (ERK) system, particularly where methods are being developed to offer a more user friendly package for use in general dental practice (Martignon 2019).

### Implications for research

Whilst a large number of studies have been conducted on the detection of early caries using visual methods, few studies were

judged to be at low risk of bias across all domains, which perhaps reflects the challenges in designing and conducting robust clinical studies in this area as discussed throughout this review. We would urge researchers to follow the STARD checklist when reporting future studies (Bossuyt 2015). In particular, any future study should ensure that participants or tooth samples are recruited consecutively or randomly and clearly state the inclusion criteria.

Studies that complete different visual scales on the same teeth or participants would be beneficial as future meta-analysis could then be conducted as a direct rather than indirect comparisons minimising confounding. This may prove burdensome for participants however, due to discomfort from multiple examinations. Other study designs that could be of interest would be diagnostic test accuracy studies which compare general dental practice routine examinations and more in depth classification systems (such as ICDAS and ERK), or randomised controlled trials that compare the use of different visual classification systems and observe the effect on disease outcomes.

Future studies should be carried out in a clinical setting, to provide a realistic assessment of performance within the oral cavity with the challenges of plaque, tooth staining, and restorations, and consider methods to minimise bias arising from the use of imperfect reference standards in clinical studies. Alternatively, *in vivo* study designs could be utilised where the visual classification systems are used on teeth due to be extracted, which permits the use of histology as a reference standard but brings challenges in terms of recruitment.

## 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 Richard Hogan and Iain Pretty for their advice during protocol development and the interpretation of findings; Associate Professor KR Ekstrand and J Bader (Emeritus Professor, UNC School of Dentistry, Chapel Hill North Carolina, USA) for their feedback on the protocol; and Derek Richards, Dr Gerd Göstemeyer (Department of Operative and Preventive Dentistry, Charité-Universitätsmedizin Berlin, Germany), Jennifer Hilgart, 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 for their support on this review.

## REFERENCES

## References to studies included in this review

**Achilleos 2013** {published data only}

Achilleos EE, Rahiotis C, Kakaboura A, Vougiouklakis G. Evaluation of a new fluorescence-based device in the detection of incipient occlusal caries lesions. *Lasers in Medical Science* 2013;**28**(1):193-201.

**Akarsu 2006** {published data only}

Akarsu S, Koprulu H. In vivo comparison of the efficacy of DIAGNOdent by visual inspection and radiographic diagnostic techniques in the diagnosis of occlusal caries. *Journal of Clinical Dentistry* 2006;**17**(3):53-8.

**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**(2):83-8.

**Bahrololoomi 2015** {published data only}

Bahrololoomi Z, Ezoddini F, Halvani N. Comparison of radiography, laser fluorescence and visual examination for diagnosing incipient occlusal caries of permanent first molars. *Journal of Dentistry / Tehran University of Medical Sciences* 2015;**12**(5):324-32.

**Bottenberg 2016** {published data only}

Bottenberg P, Jacquet W, Behrens C, Stachniss V, Jablonski-Momeni A. Comparison of occlusal caries detection using the ICDAS criteria on extracted teeth or their photographs. *BMC Oral Health* 2016;**16**(1):93.

**Braga 2009** {published data only}

Braga MM, Morais CC, Nakama RCS, Leamari VM, Siqueira WL, Mendes FM. In vitro performance of methods of approximal caries detection in primary molars. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology* 2009;**108**(4):e35-41.

**Braga 2010** {published data only}

Braga MM, Ekstrand KR, Martignon S, Imparato JC, Ricketts DN, Mendes FM. Clinical performance of two visual scoring systems in detecting and assessing activity status of occlusal caries in primary teeth. *Caries Research* 2010;**44**(3):300-8.

**Braun 2017** {published data only}

Braun A, Guiraud LM, Frankenberger R. Histological validation of ICDAS II and radiological assessment of occlusal carious lesions in permanent teeth. *Odontology* 2017;**105**(1):46-53.

**Bussaneli 2015** {published data only}

Bussaneli DG, Restrepo M, Boldieri T, Albertoni TH, Santos-Pinto L, Cordeiro RCL. Proximal caries lesion detection in primary teeth: does this justify the association of diagnostic methods? *Lasers in Medical Science* 2015;**30**(9):2239-44.

**Bussaneli 2015a** {published data only}

Bussaneli DG, Boldieri T, Diniz MB, Rivera LM, Santos-Pinto L, Cordeiro RC. Influence of professional experience on detection and treatment decision of occlusal caries lesions in primary teeth. *International Journal of Paediatric Dentistry* 2015;**25**(6):418-27.

**Campos 2017** {published data only}

Campos SAG, Vieira MLO, Sousa FB. Correlation between ICDAS and histology: differences between stereomicroscopy and microradiography with contrast solution as histological techniques. *PLOS One* 2017;**12**(8):e0183432.

**Carvalho 2018** {published data only}

Carvalho RN, Letieri ADS, Vieira TI, Santos TMPD, Lopes RT, Neves AA, et al. Accuracy of visual and image-based ICDAS criteria compared with a micro-CT gold standard for caries detection on occlusal surfaces. *Pesquisa Odontologica Brasileira [Brazilian Oral Research]* 2018;**32**:e60.

**Castilho 2016** {published data only}

Castilho LS, Cotta FV, Bueno AC, Moreira AN, Ferreira EF, Magalhaes CS. Validation of DIAGNOdent laser fluorescence and the International Caries Detection and Assessment System (ICDAS) in diagnosis of occlusal caries in permanent teeth: an in vivo study. *European Journal of Oral Sciences* 2016;**124**(2):188-94.

**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**(2):161-9.

**Cinar 2013** {published data only}

Cinar C, Atabek D, Odabas ME, Olmez A. Comparison of laser fluorescence devices for detection of caries in primary teeth. *International Dental Journal* 2013;**63**(2):97-102.

**Costa 2002** {published data only}

Costa AM, Yamaguti PM, De Paula LM, Bezerra AC. In vitro study of laser diode 655 nm diagnosis of occlusal caries. *Journal of Dentistry for Children* 2002;**69**:249-53, 233.

**Da Silva 2010** {published data only}

Da Silva PRD, Marques MM, Steagall W Jr, Mendes FM, Lascala CA. Accuracy of direct digital radiography for detecting occlusal caries in primary teeth compared with conventional radiography and visual inspection: an in vitro study. *Dentomaxillofacial Radiology* 2010;**39**(6):362-7.

**Diniz 2009** {published data only}

Diniz MB, Rodrigues JA, Hug I, Cordeiro RC, Lussi A. Reproducibility and accuracy of the ICDAS-II for occlusal caries detection. *Community Dentistry and Oral Epidemiology* 2009;**37**(5):399-404.

**Diniz 2011** {published data only}

Diniz MB, Lima LM, Eckert G, Zandona AG, Cordeiro RC, Pinto LS. In vitro evaluation of ICDAS and radiographic examination

of occlusal surfaces and their association with treatment decisions. *Operative Dentistry* 2011;**36**(2):133-42.

**Diniz 2012** {published data only}

Diniz MB, Boldieri T, Rodrigues JA, Santos-Pinto L, Lussi A, Cordeiro RC. The performance of conventional and fluorescence-based methods for occlusal caries detection: an in vivo study with histologic validation. *Journal of the American Dental Association* 2012;**143**(4):339-50.

**Diniz 2019** {published data only}

Diniz MB, Campos PH, Wilde S, Cordeiro RCL, Zandona AGF. Performance of light-emitting diode device in detecting occlusal caries in the primary molars. *Lasers in Medical Science* 2019;**34**(6):1235-41.

**Ekstrand 2007** {published data only}

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

**Ekstrand 2011** {published data only}

Ekstrand KR, Luna LE, Promisiero L, Cortes A, Cuevas S, Reyes JF, et al. The reliability and accuracy of two methods for proximal caries detection and depth on directly visible proximal surfaces: an in vitro study. *Caries Research* 2011;**45**(2):93-9.

**Ekstrand 2011a** {published data only}

Ekstrand KR, Luna LE, Promisiero L, Cortes A, Cuevas S, Reyes JF, et al. The reliability and accuracy of two methods for proximal caries detection and depth on directly visible proximal surfaces: an in vitro study. *Caries Research* 2011;**45**(2):93-9.

**Erten 2005** {published data only}

Erten H, Uctasli MB, Akarlan ZZ, Uzun O, Baspinar E. The assessment of unaided visual examination, intraoral camera and operating microscope for the detection of occlusal caries lesions. *Operative Dentistry* 2005;**30**(2):190-4.

**Forge 2003** {published data only}

Forge AH, Pine CM, Pitts NB. The assessment of an intra-oral video camera as an aid to occlusal caries detection. *International Dental Journal* 2003;**53**(1):3-6.

**Freitas 2016** {published data only}

Freitas LA, Santos MT, Guaré RO, Lussi A, Diniz MB. Association between visual inspection, caries activity status, and radiography with treatment decisions on approximal caries in primary molars. *Pediatric Dentistry* 2016;**38**(2):140-7.

**Goel 2009** {published data only}

Goel A, Chawla HS, Gauba K, Goyal A. Comparison of validity of DIAGNOdent with conventional methods for detection of occlusal caries in primary molars using the histological gold standard: an in vivo study. *Journal of the Indian Society of Pedodontics and Preventive Dentistry* 2009;**27**(4):227-34.

**Haak 2002** {published data only}

Haak R, Wicht MJ, Hellmich M, Gossmann A, Noack MJ. The validity of proximal caries detection using magnifying visual aids. *Caries Research* 2002;**36**(4):249-55.

**Hintze 2003** {published data only}

Hintze H, Wenzel A. Diagnostic outcome of methods frequently used for caries validation. A comparison of clinical examination, radiography and histology following hemisectioning and serial tooth sectioning. *Caries Research* 2003;**37**(2):115-24.

**Hintze 2003a** {published data only}

Hintze H, Wenzel A. Diagnostic outcome of methods frequently used for caries validation. A comparison of clinical examination, radiography and histology following hemisectioning and serial tooth sectioning. *Caries Research* 2003;**37**(2):115-24.

**Huth 2010** {published data only}

Huth KC, Lussi A, Gyax M, Thum M, Crispin A, Paschos E, et al. In vivo performance of a laser fluorescence device for the approximal detection of caries in permanent molars. *Journal of Dentistry* 2010;**38**(12):1019-26.

**Iranzo-Cortes 2017** {published data only}

Iranzo-Cortes JE, Terzic S, Montiel-Company JM, Almerich-Silla JM. Diagnostic validity of ICDAS and DIAGNOdent combined: an in vitro study in pre-cavitated lesions. *Lasers in Medical Science* 2017;**32**(3):543-8.

**Iranzo-Cortes 2018** {published data only}

Iranzo-Cortes JE, Almarche-Tarazona M, Montiel-Company JM, Almerich-Silla JM. Diagnostic validity of ICDAS II, VistaProof and a combination of these two methods. An in vitro study in pre-cavitated lesions. *Lasers in Surgery and Medicine* 2018;**50**(2):166-73.

**Jablonski-Momeni 2008** {published data only}

Jablonski-Momeni A, Stachniss V, Ricketts DN, Heinzel-Gutenbrunner M, Pieper K. Reproducibility and accuracy of the ICDAS-II for detection of occlusal caries in vitro. *Caries Research* 2008;**42**(2):79-87.

**Jablonski-Momeni 2012** {published data only}

Jablonski-Momeni A, Stucke J, Steinberg T, Heinzel-Gutenbrunner M. Use of ICDAS-II, fluorescence-based methods, and radiography in detection and treatment decision of occlusal caries lesions: an in vitro study. *International Journal of Dentistry* 2012;**2012**:371595.

**Kim 2017** {published data only}

Kim ES, Lee ES, Kang SM, Jung EH, de Josselin de Jong E, Jung HI, et al. A new screening method to detect proximal dental caries using fluorescence imaging. *Photodiagnosis and Photodynamic Therapy* 2017;**20**:257-62.

**Ko 2015** {published data only}

Ko HY, Kang SM, Kim HE, Kwon HK, Kim BI. Validation of quantitative light-induced fluorescence-digital (QLF-D) for the detection of approximal caries in vitro. *Journal of Dentistry* 2015;**43**(5):568-75.

**Kockanat 2017** {published data only}

Kockanat A, Unal M. In vivo and in vitro comparison of ICDAS II, DIAGNOdent pen, CarieScan PRO and SoproLife camera for occlusal caries detection in primary molar teeth. *European Journal of Paediatric Dentistry* 2017;**18**(2):99-104.

**Kucukyilmaz 2015** {published data only}

Kucukyilmaz E, Sener Y, Botsali MS. In vivo and in vitro performance of conventional methods, DIAGNOdent, and an electronic caries monitor for occlusal caries detection in primary teeth. *Pediatric Dentistry* 2015;**37**(4):E14-22.

**Kuhnisch 2009** {published data only}

Kuhnisch J, Ifland S, Tranaeus S, Heinrich-Weltzien R. Comparison of visual inspection and different radiographic methods for dentin caries detection on occlusal surfaces. *Dentomaxillofacial Radiology* 2009;**38**(7):452-7.

**Kuhnisch 2009a** {published data only}

Kuhnisch J, Goddon I, Berger S, Senkel H, Bucher K, Oehme T, et al. Development, methodology and potential of the new Universal Visual Scoring System (UniViSS) for caries detection and diagnosis. *International Journal of Environmental Research and Public Health* 2009;**6**(9):2500-9.

**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.

**Mendes 2006** {published data only}

Mendes FM, Ganzerla E, Nunes AF, Puig AV, Imparato JC. Use of high-powered magnification to detect occlusal caries in primary teeth. *American Journal of Dentistry* 2006;**19**(1):19-22.

**Mitropoulos 2010** {published data only}

Mitropoulos P, Rahiotis C, Stamatakis H, Kakaboura A. Diagnostic performance of the visual caries classification system ICDAS II versus radiography and micro-computed tomography for proximal caries detection: an in vitro study. *Journal of Dentistry* 2010;**38**(11):859-67.

**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.

**Neuhaus 2011** {published data only}

Neuhaus KW, Rodrigues JA, Hug I, Stich H, Lussi A. Performance of laser fluorescence devices, visual and radiographic examination for the detection of occlusal caries in primary molars. *Clinical Oral Investigations* 2011;**15**(5):635-41.

**Neuhaus 2015** {published data only}

Neuhaus KW, Jost F, Perrin P, Lussi A. Impact of different magnification levels on visual caries detection with ICDAS. *Journal of Dentistry* 2015;**43**(12):1559-64.

**Novaes 2009** {published data only}

Novaes TF, Matos R, Braga MM, Imparato JC, Raggio DP, Mendes FM. Performance of a pen-type laser fluorescence device and conventional methods in detecting approximal caries lesions in primary teeth - in vivo study. *Caries Research* 2009;**43**(1):36-42.

**Novaes 2010** {published data only}

Novaes TF, Matos R, Raggio DP, Imparato JC, Braga MM, Mendes FM. Influence of the discomfort reported by children on the performance of approximal caries detection methods. *Caries Research* 2010;**44**(5):465-71.

**Novaes 2012** {published data only}

Novaes TF, Matos R, Gimenez T, Braga MM, Benedetto DE, Mendes FM. Performance of fluorescence-based and conventional methods of occlusal caries detection in primary molars - an in vitro study. *International Journal of Paediatric Dentistry* 2012;**22**(6):459-66.

**Paula 2011** {published data only}

De Paula AB, Campos JADB, Diniz MB, Hebling J, Rodrigues JA. In situ and in vitro comparison of laser fluorescence with visual inspection in detecting occlusal caries lesions. *Lasers in Medical Science* 2011;**26**:1-5.

**Pereira 2011** {published data only}

Pereira AC, Eggertsson H, González-Cabezas C, Zero DT, Eckert GJ, Mialhe FL. Quantitative light-induced fluorescence (QLF) in relation to other technologies and conventional methods for detecting occlusal caries in permanent teeth. *Brazilian Journal of Oral Sciences* 2011;**10**(1):27-32.

**Qudeimat 2019** {published data only}

Qudeimat MA, Altarakemah Y, Alomari Q, Alshawaf N, Honkala E. The impact of ICDAS on occlusal caries treatment recommendations for high caries risk patients: an in vitro study. *BMC Oral Health* 2019;**19**(1):41.

**Rocha 2003** {published data only}

Rocha RO, Ardenghi TM, Oliveira LB, Rodrigues CR, Ciamponi AL. In vivo effectiveness of laser fluorescence compared to visual inspection and radiography for the detection of occlusal caries in primary teeth. *Caries Research* 2003;**37**(6):437-41.

**Rodrigues 2008** {published data only}

Rodrigues JA, Hug I, Diniz MB, Lussi A. Performance of fluorescence methods, radiographic examination and ICDAS II on occlusal surfaces in vitro. *Caries Research* 2008;**42**(4):297-304.

**Rodrigues 2009** {published data only}

Rodrigues JA, Diniz MB, Josgrillberg EB, Cordeiro RC. In vitro comparison of laser fluorescence performance with visual examination for detection of occlusal caries in permanent and primary molars. *Lasers in Medical Science* 2009;**24**(4):501-6.

**Rodrigues 2009a** {published data only}

Rodrigues JA, Diniz MB, Josgrillberg EB, Cordeiro RC. In vitro comparison of laser fluorescence performance with visual

examination for detection of occlusal caries in permanent and primary molars. *Lasers in Medical Science* 2009;**24**(4):501-6.

**Rodrigues 2013** {published data only}

Rodrigues JA, Oliveira RS de, Hug I, Neuhaus K, Lussi A. Performance of experienced dentists in Switzerland after an e-learning program on ICDAS occlusal caries detection. *Journal of Dental Education* 2013;**77**(8):1086-91.

**Seremidi 2012** {published data only}

Seremidi K, Lagouvardos P, Kavvadia K. Comparative in vitro validation of VistaProof and DIAGNOdent pen for occlusal caries detection in permanent teeth. *Operative Dentistry* 2012;**37**(3):234-45.

**Shi 2000** {published data only}

Shi XQ, Welander U, Angmar-Mansson B. Occlusal caries detection with KaVo DIAGNOdent and radiography: an in vitro comparison. *Caries Research* 2000;**34**:151-8.

**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.

**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.

**Souza 2013** {published data only}

Souza JF, Boldieri T, Diniz MB, Rodrigues JA, Lussi A, Cordeiro RC. Traditional and novel methods for occlusal caries detection: performance on primary teeth. *Lasers in Medical Science* 2013;**28**:287-95.

**Soviero 2012** {published data only}

Soviero VM, Leal SC, Silva RC, Azevedo RB. Validity of MicroCT for in vitro detection of proximal carious lesions in primary molars. *Journal of Dentistry* 2012;**40**(1):35-40.

**Sridhar 2009** {published data only}

Sridhar N, Tandon S, Rao N. A comparative evaluation of DIAGNOdent with visual and radiography for detection of occlusal caries: an in vitro study. *Indian Journal of Dental Research* 2009;**20**:326-31.

**Teo 2014** {published data only}

Teo TK, Ashley PF, Louca C. An in vivo and in vitro investigation of the use of ICDAS, DIAGNOdent pen and CarieScan PRO for the detection and assessment of occlusal caries in primary molar teeth. *Clinical Oral Investigations* 2014;**18**(3):737-44.

**Tonkaboni 2019** {published data only}

Tonkaboni A, Saffarpour A, Aghapourzangeneh F, Fard MJK. Comparison of diagnostic effects of infrared imaging and bitewing radiography in proximal caries of permanent teeth. *Lasers in Medical Science* 2019;**34**(5):873-9.

**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.

## References to studies excluded from this review

**Abrams 2017** {published data only}

Abrams SH, Sivagurunathan KS, Silvertown JD, Wong B, Hellen A, Mandelis A, et al. Correlation with caries lesion depth of the Canary System, DIAGNOdent and ICDAS II. *Open Dentistry Journal* 2017;**11**:679-89.

**Almosa 2014** {published data only}

Almosa NA, Lundgren T, Aldrees AM, Birkhed D, Kjellberg H. Diagnosing the severity of buccal caries lesions in governmental and private orthodontic patients at debonding, using the ICDAS-II and the DIAGNOdent Pen. *Angle Orthodontist* 2014;**84**(3):430-6.

**Amaechi 2013** {published data only}

Amaechi BT, Chedjieu I, Lozano-Pineda J. Clinical evaluation of an enhanced white light and fluorescence device for early detection of caries lesions. *Journal of Clinical Dentistry* 2013;**24**(2):43-8.

**Apostolopoulou 2009** {published data only}

Apostolopoulou D, Lagouvardos P, Kavvadia K, Papagiannoulis L. Histological validation of a laser fluorescence device for occlusal caries detection in primary molars. *European Archives of Paediatric Dentistry* 2009;**10** Suppl 1:11-5.

**Ashley 2000** {published data only}

Ashley P. Diagnosis of occlusal caries in primary teeth. *International Journal of Paediatric Dentistry* 2000;**10**(2):166-71.

**Askaroglou 2011** {published data only}

Askaroglou E, Kavvadia K, Lagouvardos P, Papagiannoulis L. Effect of sealants on laser fluorescence caries detection in primary teeth. *Lasers in Medical Science* 2011;**26**(1):29-34.

**Attrill 2001** {published data only}

Attrill DC, Ashley PF. Occlusal caries detection in primary teeth: a comparison of DIAGNOdent with conventional methods. *British Dental Journal* 2001;**190**:440-3.

**Bengtson 2005** {published data only}

Bengtson AL, Gomes AC, Mendes FM, Cichello LR, Bengtson NG, Pinheiro SL. Influence of examiner's clinical experience in detecting occlusal caries lesions in primary teeth. *Pediatric Dentistry* 2005;**27**:238-43.

**Bizhang 2016** {published data only}

Bizhang M, Wollenweber N, Singh-Husgen P, Danesh G, Zimmer S. Pen-type laser fluorescence device versus bitewing radiographs for caries detection on approximal surfaces. *Head & Face Medicine* 2016;**12**(1):30.

**Bozdemir 2013** {published data only}

Bozdemir E, Karaarslan ES, Ozsevik AS, Ata Cebe M, Aktan AM. In vivo performance of two devices for occlusal caries detection. *Photomedicine and Laser Surgery* 2013;**31**(7):322-7.

**Braga 2009a** {published data only}

Braga MM, Mendes FM, Martignon S, Ricketts DN, Ekstrand KR. In vitro comparison of Nyvad's system and ICDAS-II with Lesion Activity Assessment for evaluation of severity and activity of occlusal caries lesions in primary teeth. *Caries Research* 2009;**43**(5):405-12.

**Chen 2012** {published data only}

Chen J, Qin M, Ma W, Ge L. A clinical study of a laser fluorescence device for the detection of approximal caries in primary molars. *International Journal of Paediatric Dentistry* 2012;**22**(2):132-8.

**Chong 2003** {published data only}

Chong MJ, Seow WK, Purdie DM, Cheng E, Wan V. Visual-tactile examination compared with conventional radiography, digital radiography, and Diagnodent in the diagnosis of occlusal occult caries in extracted premolars. *Pediatric Dentistry* 2003;**25**(4):341-9.

**Chu 2010** {published data only}

Chu CH, Lo ECM, You DSH. Clinical diagnosis of fissure caries with conventional and laser-induced fluorescence techniques. *Lasers in Medical Science* 2010;**25**(3):355-62.

**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**(1):8-16.

**Costa 2007** {published data only}

Costa AM, Bezzerra AC, Fuks AB. Assessment of the accuracy of visual examination, bite-wing radiographs and DIAGNOdent on the diagnosis of occlusal caries. *European Archives of Paediatric Dentistry* 2007;**8**:118-22.

**Diniz 2011a** {published data only}

Diniz MB, Sciasci P, Rodrigues JA, Lussi A, Cordeiro RCL. Influence of different professional prophylactic methods on fluorescence measurements for detection of occlusal caries. *Caries Research* 2011;**45**(3):264-8.

**Dong 2007** {published data only}

Dong C, Choo-Smith L, Werner J, Ko A, Hewko M, Cleghorn B, et al. Comparing clinical and spectroscopic methods for detecting early dental caries. *Journal of Dental Research* 2007;**85**(Spec Iss B):1241.

**Duruturk 2011** {published data only}

Duruturk L, Ciftci A, Baharoglu S, Oztuna D. Clinical evaluation of DIAGNOdent in detection of occlusal caries in newly erupted noncavitated first permanent molars in caries-active children. *Operative Dentistry* 2011;**36**:348-55.

**El-Damanhoury 2014** {published data only}

El-Damanhoury HM, Fakhruddin KS, Awad MA. Effectiveness of teaching International Caries Detection and Assessment System II and its e-learning program to freshman dental students on occlusal caries detection. *European Journal of Dentistry* 2014;**8**(4):493-7.

**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.

**Erten 2006** {published data only}

Erten H, Uctasli MB, Akarslan ZZ, Uzun O, Semiz M. Restorative treatment decision making with unaided visual examination, intraoral camera and operating microscope. *Operative Dentistry* 2006;**31**(1):55-9.

**Forgie 2002** {published data only}

Forgie AH, Pine CM, Pitts NB. The use of magnification in a preventive approach to caries detection. *Quintessence International* 2002;**33**(1):13-6.

**Fyffe 2000** {published data only}

Fyffe HE, Deery C, Nugent ZJ, Nuttall NM, Pitts NB. In vitro validity of the Dundee Selectable Threshold Method for caries diagnosis (DSTM). *Community Dentistry and Oral Epidemiology* 2000;**28**(1):52-8.

**Goel 2016** {published data only}

Goel D, Sandhu M, Jhingan P, Sachdev V. Effectiveness of air drying and magnification methods for detecting initial caries on occlusal surfaces using three different diagnostic aids. *Journal of Clinical Pediatric Dentistry* 2016;**40**(3):221-6.

**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.

**Graye 2012** {published data only}

Graye M, Markowitz K, Strickland M, Guzy G, Burke M, Houpt M. In vitro evaluation of the Spectra™ early caries detection system. *Journal of Clinical Dentistry* 2012;**23**(1):1.

**Heinrich-Weltzien 2003** {published data only}

Heinrich-Weltzien R, Kuhnisch J, Oehme T, Ziehe A, Stosser L, Garcia-Godoy F. Comparison of different DIAGNOdent cut-off limits for in vivo detection of occlusal caries. *Operative Dentistry* 2003;**28**(6):672-80.

**Huysmans 1998** {published data only}

Huysmans MC, Longbottom C, Pitts N. Electrical methods in occlusal caries diagnosis: an in vitro comparison with visual inspection and bite-wing radiography. *Caries Research* 1998;**32**(5):324-9.

**Iranzo-Cortés 2018** {published data only}

Iranzo-Cortés JE, Almarche-Tarazona M, Montiel-Company JM, Almerich-Silla JM. Diagnostic validity of ICDAS II, VistaProof

and a combination of these two methods. An in vitro study in pre-cavitated lesions. *Lasers in Surgery and Medicine* 2018;**50**(2):166-73.

**Jablonski-Momeni 2009** {published data only}

Jablonski-Momeni A, Ricketts DN, Heinzl-Gutenbrunner M, Stoll R, Stachniss V, Pieper K. Impact of scoring single or multiple occlusal lesions on estimates of diagnostic accuracy of the visual ICDAS-II system. *International Journal of Dentistry* 2009;**2009**:798283.

**Jablonski-Momeni 2013** {published data only}

Jablonski-Momeni A, Liebegall F, Stoll R, Heinzl-Gutenbrunner M, Pieper K. Performance of a new fluorescence camera for detection of occlusal caries in vitro. *Lasers in Medical Science* 2013;**28**(1):101-9.

**Jablonski-Momeni 2018** {published data only}

Jablonski-Momeni A, Moos J, Manesh VS, Stoll R. Diagnostic accuracy of a bioluminescence system for the assessment of caries activity on occlusal surfaces. *Caries Research* 2018;**52**(4):279-87.

**Kavvadia 2008** {published data only}

Kavvadia K, Lagouvardos P. Clinical performance of a diode laser fluorescence device for the detection of occlusal caries in primary teeth. *International Journal of Paediatric Dentistry* 2008;**18**(3):197-204.

**Kavvadia 2012** {published data only}

Kavvadia K, Lagouvardos P, Apostolopoulou D. Combined validity of DIAGNOdent™ and visual examination for in vitro detection of occlusal caries in primary molars. *Lasers in Medical Science* 2012;**27**(2):313-9.

**Khalaf 2018** {published data only}

Khalaf ME, Qudeimat M, Alomari QD, Al-Tarakmeh Y, Honkala E. Impact of magnifying loupes on interexaminer agreement in the detection of noncavitated occlusal carious lesions. *Oral Health & Preventive Dentistry* 2018;**16**(4):375-80.

**Kidd 2003** {published data only}

Kidd EAM, Banerjee A, Ferrier S, Longbottom C, Nugent Z. Relationships between a clinical-visual scoring system and two histological techniques: a laboratory study on occlusal and approximal carious lesions. *Caries Research* 2003;**37**(2):125-9.

**Kordic 2003** {published data only}

Kordic A, Lussi A, Luder HU. Performance of visual inspection, electrical conductance and laser fluorescence in detecting occlusal caries in vitro. *Schweizer Monatsschrift für Zahnmedizin SMfZ* 2003;**113**(8):852-9.

**Kuhnisch 2011** {published data only}

Kuhnisch J, Bucher K, Henschel V, Albrecht A, Garcia-Godoy F, Mansmann U, et al. Diagnostic performance of the universal visual scoring system (UniVISS) on occlusal surfaces. *Clinical Oral Investigations* 2011;**15**(2):215-23.

**Lussi 1991** {published data only}

Lussi A. Validity of diagnostic and treatment decisions of fissure caries. *Caries Research* 1991;**25**(4):296-303.

**Lussi 2003** {published data only}

Lussi A, Francescut P. Performance of conventional and new methods for the detection of occlusal caries in deciduous teeth. *Caries Research* 2003;**37**(1):2-7.

**Melo 2017** {published data only}

Melo M, Pascual A, Camps I, Campo A Del, Ata-Ali J. Caries diagnosis using light fluorescence devices in comparison with traditional visual and tactile evaluation: a prospective study in 152 patients. *Odontology* 2017;**105**(3):283-90.

**Mialhe 2011** {published data only}

Mialhe FL, Ambrosano GM. Validity of caries-detection methods under epidemiological setting. *American Journal of Dentistry* 2011;**24**(6):363-6.

**Mitropoulos 2012** {published data only}

Mitropoulos P, Rahiotis C, Kakaboura A, Vougiouklakis G. The impact of magnification on occlusal caries diagnosis with implementation of the ICDAS II criteria. *Caries Research* 2012;**46**(1):82-6.

**Neuhaus 2015a** {published data only}

Neuhaus KW, Ciucchi P, Rodrigues JA, Hug I, Emerich M, Lussi A. Diagnostic performance of a new red light LED device for approximal caries detection. *Lasers in Medical Science* 2015;**30**(5):1443-7.

**Novaes 2012a** {published data only}

Novaes TF, Matos R, Celiberti P, Braga MM, Mendes FM. The influence of interdental spacing on the detection of proximal caries lesions in primary teeth. *Brazilian Oral Research* 2012;**26**(4):293-9.

**Pereira 2009** {published data only}

Pereira AC, Eggertsson H, Martinez-Mier EA, Mialhe FL, Eckert GJ, Zero DT. Validity of caries detection on occlusal surfaces and treatment decisions based on results from multiple caries detection methods. *European Journal of Oral Sciences* 2009;**117**(1):51-7.

**Piovesan 2013** {published data only}

Piovesan C, Moro BL, Lara JS, Ardenghi TM, Guedes RS, Haddad AE, et al. Laboratorial training of examiners for using a visual caries detection system in epidemiological surveys. *BMC Oral Health* 2013;**13**(1):49.

**Qudeimat 2016** {published data only}

Qudeimat MA, Alomari QD, Altarakemah Y, Alshawaf N, Honkala EJ. Variables affecting the inter- and intra-examiner reliability of ICDAS for occlusal caries diagnosis in permanent molars. *Journal of Public Health Dentistry* 2016;**76**(1):9-16.

**Reis 2006** {published data only}

Reis A, Mendes FM, Angnes V, Angnes G, Grande RH, Loguercio AD. Performance of methods of occlusal caries

detection in permanent teeth under clinical and laboratory conditions. *Journal of Dentistry* 2006;**34**(2):89-96.

**Ricketts 1995** {published data only}

Ricketts DNJ, Kidd EAM, Smith BGN, Wilson RF. Clinical and radiographic diagnosis of occlusal caries: a study in vitro. *Journal of Oral Rehabilitation* 1995;**22**(1):15-20.

**Rodrigues 2008a** {published data only}

Rodrigues Jde A, de Vita TM, Cordeiro Rde C. In vitro evaluation of the influence of air abrasion on detection of occlusal caries lesions in primary teeth. *Pediatric Dentistry* 2008;**30**(1):15-8.

**Senel 2010** {published data only}

Senel B, Kamburoglu K, Ucok O, Yuksel SP, Ozen T, Avsever H. Diagnostic accuracy of different imaging modalities in detection of proximal caries. *DentoMaxilloFacial Radiology* 2010;**39**(8):501-11.

**Shoaib 2009** {published data only}

Shoaib L, Deery CM, Ricketts DNJ, Nugent ZJ. Validity and reproducibility of ICDAS II in primary teeth. *Caries Research* 2009;**43**(6):442-8.

**Silva 2008** {published data only}

da Silva Neto JM, dos Santos RL, Sampaio MC, Sampaio FC, Passos IA. Radiographic diagnosis of incipient proximal caries: an ex-vivo study. *Brazilian Dental Journal* 2008;**19**(2):97-102.

**Subka 2019** {published data only}

Subka S, Rodd H, Nugent Z, Deery C. In vivo validity of proximal caries detection in primary teeth, with histological validation. *International Journal of Paediatric Dentistry* 2019;**29**(4):429-38.

**Umemori 2010** {published data only}

Umemori S, Tonami K, Nitta H, Matakai S, Araki K. The possibility of digital imaging in the diagnosis of occlusal caries. *International Journal of Dentistry* 2010;**2010**:860515.

**Ünal 2019** {published data only}

Ünal M, Koçkanat A, Güler S, Gültürk E. Diagnostic performance of different methods in detecting incipient non-cavitated occlusal caries lesions in permanent teeth. *Journal of Clinical Pediatric Dentistry* 2019;**43**(3):173-9.

**Virajsilp 2005** {published data only}

Virajsilp V, Thearmontree A, Aryatawong S, Paiboonwarachat D. Comparison of proximal caries detection in primary teeth between laser fluorescence and bitewing radiography. *Pediatric Dentistry* 2005;**27**(6):493-9.

**Wenzel 1990** {published data only}

Wenzel A, Fejerskov O, Kidd E, Joyston-Bechal S, Groeneveld A. Depth of occlusal caries assessed clinically, by conventional film radiographs, and by digitized, processed radiographs. *Caries Research* 1990;**24**(5):327-33.

**White 1978** {published data only}

White GE, Tsamtouris A, Williams DL. Early detection of occlusal caries by measuring the electrical resistance of the tooth. *Journal of Dental Research* 1978;**57**(2):195-200.

**Zandona 2009** {published data only}

Zandona AGF, Al-Shiha S, Eggertsson H, Eckert G. Student versus faculty performance using a new visual criteria for the detection of caries on occlusal surfaces: an in vitro examination with histological validation. *Operative Dentistry* 2009;**34**(5):598-604.

## References to studies awaiting assessment

**Luczaj-Cepowicz 2019** {published data only}

Luczaj-Cepowicz E, Marczuk-Kolada G, Obidzinska M, Sidun J. Diagnostic validity of the use of ICDAS II and DIAGNOdent pen verified by micro-computed tomography for the detection of occlusal caries lesions - an in vitro evaluation. *Lasers in Medical Science* 2019;**34**(8):1655-63.

## Additional references

**Adult Dental Health Survey 2009**

NHS Digital. Adult Dental Health Survey 2009 - Summary report and thematic series. [digital.nhs.uk/data-and-information/publications/statistical/adult-dental-health-survey/adult-dental-health-survey-2009-summary-report-and-thematic-series#resources](https://digital.nhs.uk/data-and-information/publications/statistical/adult-dental-health-survey/adult-dental-health-survey-2009-summary-report-and-thematic-series#resources) (accessed prior to 6 July 2020).

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

**Atkins 2004**

Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, et al. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches The GRADE Working Group. *BMC Health Services Research* 2004;**4**(1):38.

**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 2015**

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;**351**:h552777.

**Braga 2009b**

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.

**Chu 2006**

Chu H, Cole SR. Bivariate meta-analysis for sensitivity and specificity with sparse data: a generalized linear mixed model approach (letter to the Editor). *Journal of Clinical Epidemiology* 2006;**59**(12):1331-2.

**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 [srda.cochrane.org](http://srda.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 [srda.cochrane.org](http://srda.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.

**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.

**Fee 2020**

Fee PA, Macey R, Walsh T, Clarkson JE, Ricketts D. Tests to detect and inform the diagnosis of root caries. *Cochrane Database of Systematic Reviews* 2020, Issue 12. Art. No: CD013806. [DOI: [10.1002/14651858.CD013806](https://doi.org/10.1002/14651858.CD013806)]

**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.

**Fyffe 2000a**

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.

**Gimenez 2015**

Gimenez T, Piovesan C, Braga MM, Raggio DP, Deery C, Ricketts DN, et al. Visual inspection for caries detection: a systematic review and meta-analysis. *Journal of Dental Research* 2015;**94**(7):895-904.

**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.

**Harbord 2009**

Harbord RM, Whiting P. metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata Journal* 2009;**9**(2):211-29.

**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.

**Hsu 2011**

Hsu J, Brożek JL, Terracciano L, Kreis J, Compalati E, Stein AT, et al. Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. *Implementation Science* 2011;**6**(1):62.

**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.

**Leeflang 2008**

Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM. 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](http://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 A-M, 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)]

**Macey 2021**

Macey R, Walsh T, Riley P, Glenny A-M, Worthington HV, Clarkson JE, et al. Electrical conductance for the detection of dental caries. *Cochrane Database of Systematic Reviews* 2021, Issue 3. Art. No: CD014547. [DOI: [10.1002/14651858.CD014547](https://doi.org/10.1002/14651858.CD014547)]

**Macey 2021a**

Macey R, Walsh T, Riley P, Hogan R, Glenny A-M, Worthington HV, et al. 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)]

**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, Thombs 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.

**Nyvad 1999**

Nyvad B, Machiulskiene V, Bælum V. Reliability of a new caries diagnostic system differentiating between active and inactive caries lesions. *Caries Research* 1999;**33**:252-60.

**Ormond 2010**

Ormond C, Douglas G, Pitts N. The use of the International Caries Detection and Assessment System (ICDAS) in a National Health Service general dental practice as part of an oral health assessment. *Primary Dental Care* 2010;**17**(4):153-60.

**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 (August 2016). Available at [methods.cochrane.org/sdt/](http://methods.cochrane.org/sdt/).

**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 1997**

Pitts NB, Evans DJ, Pine CM. British Association for the Study of Community Dentistry (BASCD) diagnostic criteria for caries prevalence surveys-1996/97. *Community Dental Health* 1997;**14**:6-9.

**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 children, 2014-15. National protocol. [www.nwph.net/dentalhealth/14\\_15\\_5yearold/Protocol\\_2014\\_15\\_5%20yr%20olds%20v2.pdf](http://www.nwph.net/dentalhealth/14_15_5yearold/Protocol_2014_15_5%20yr%20olds%20v2.pdf) (accessed prior to 6 July 2020).

**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.

**Schünemann 2020**

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.

**Stata 14 [Computer program]**

Stata. Version 15. College Station, TX, USA: StataCorp, 2017. Available at [www.stata.com](http://www.stata.com).

**Steele 2011**

Steele J, O'Sullivan I. Executive Summary: Adult Dental Health Survey 2009. Available at [files.digital.nhs.uk/publicationimport/pub01xxx/pub01086/adul-dent-heal-surv-summ-them-exec-2009-rep2.pdf](http://files.digital.nhs.uk/publicationimport/pub01xxx/pub01086/adul-dent-heal-surv-summ-them-exec-2009-rep2.pdf).

**Takwoingi 2016**

Takwoingi Y. Meta-analysis of test accuracy studies in Stata: a bivariate model approach. Version 1.1. [methods.cochrane.org/sdt/](http://methods.cochrane.org/sdt/) 2016.

**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.

**Walsh 2021**

Walsh T, Macey R, Riley P, Glenny A-M, Schwendicke F, Worthington HV, et al. Imaging modalities to inform the detection and diagnosis of early caries. *Cochrane Database of Systematic Reviews* 2021, Issue 3. Art. No: CD014545. [DOI: [10.1002/14651858.CD014545](https://doi.org/10.1002/14651858.CD014545)]

**Wenzel 2006**

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.

**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](https://apps.who.int/iris/bitstream/handle/10665/259413/WHO-NMH-NHD-17.12-eng.pdf?sequence=1) (accessed prior to 6 July 2020).

**Zandona 2019**

Zandona AF, Longbottom C, editor(s). Detection and Assessment of Dental Caries. Springer, 2019.

**Zhang 2019**

Zhang Y, Coello PA, Guyatt GH, Yepes-Nuñez JJ, Akl EA, Hazlewood G, et al. GRADE guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences - inconsistency, imprecision, and other domains. *Journal of Clinical Epidemiology* 2019;**111**:83-93.

## CHARACTERISTICS OF STUDIES

### Characteristics of included studies [ordered by study ID]

#### Achilleos 2013

##### Study characteristics

Patient Sampling	Method of sampling: selected Included conditions: no cavitation Teeth: permanent molars and premolars Surface: occlusal Restorations: excluded Sealants: unclear
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Greece Setting: extracted for orthodontic purposes Number of participants/teeth/sites: 38 teeth Prevalence: enamel = 0.95, dentine = 0.39

**Achilleos 2013** (Continued)

Index tests	<p>Category of test: visual ICDAS</p> <p>Sequence of test(s): visual, followed by VistaProof/DIAGNOdent then reference standard</p> <p>Examiner training and calibration: experienced, trained and calibrated dentists: "Two dentists with 1-year clinical experience in the ICDAS criteria"</p> <p>Teeth cleaning prior to examination: calculus and debris were removed by paste and brush burr</p> <p>Tooth drying prior to examination: yes, 5 seconds</p> <p>Threshold applied: ICDAS (0): sound tooth surface, (1): first visual change in dry enamel, (2): distinct visual change in moist enamel, (3): localised enamel breakdown due to caries with no visible dentine or underlying shadow, (4): underlying dark shadow in dentine with or without localised enamel breakdown, (5): distinct cavity with visible dentine, and (6): extensive distinct cavity with visible dentine</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: following index test</p> <p>Training of examiner: experienced, same examiner as index test</p> <p>Blinding to index test: no</p> <p>Multiple tests: no</p> <p>Site selection: 3 sections</p> <p>Target condition: caries free, early enamel, deep enamel, outer dentine, dentine, deep 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	Multiple examiners reported so examiner 1's values reported

**Methodological quality**

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

**Achilleos 2013** (Continued)

<b>Could the selection of patients have introduced bias?</b>	High risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>	High
<b>DOMAIN 2: Index Test (All tests)</b>	
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)?	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	High
<b>DOMAIN 3: Reference Standard</b>	
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
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Akarsu 2006**
**Study characteristics**

Patient Sampling	Method of sampling: unclear Included conditions: unclear, "suspected to have occlusal caries" but unclear to what level
------------------	--

**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

**Akarsu 2006** (Continued)

	<p>Teeth: permanent molars (third molars excluded)</p> <p>Surface: occlusal</p> <p>Restorations: excluded</p> <p>Sealants: no</p>
Patient characteristics and setting	<p>Age: 18 to 25 years</p> <p>Sex: 87 female, 74 male</p> <p>Ethnicity: not reported</p> <p>Country: Turkey</p> <p>Setting: restorative clinic at dental hospital</p> <p>Number of participants/teeth/sites: 161 participants, 187 teeth</p> <p>Prevalence: enamel = 0.77, dentine = 0.52</p>
Index tests	<p>Category of test: visual - "a slightly modified version of the system previously reported by <a href="#">Ekstrand 1997</a>"</p> <p>Sequence of test(s): visual, radiograph, DIAGNOdent, then reference standard (visual, radiograph and DIAGNOdent used as part of reference standard) - assumed that visual assessment was performed before other tests</p> <p>Examiner training and calibration: unclear</p> <p>Teeth cleaning prior to examination: calculus and plaque removed using a scaler and rubber cup - no pumice used</p> <p>Tooth drying prior to examination: 8 seconds</p> <p>Threshold applied: no or slight change, opacity visible after air drying, discolouration visible after drying, opacity visible, discolouration visible, localised enamel breakdown, enamel cavitation exposing dentine</p>
Target condition and reference standard(s)	<p>Category: used a hybrid visual, radiograph, and a fluorescence test; to determine excavation as appropriate, decay was "removed by using rotational cutting devices" and the cavities assessed visually, i.e. excavation</p> <p>Sequence of index test and reference standard: following index test</p> <p>Training of examiner: experienced, same examiner as index test</p> <p>Blinding to index test: no</p> <p>Multiple tests: no</p> <p>Site selection: 3 sections</p> <p>Target condition: caries free, early enamel, deep enamel, outer dentine, dentine, deep 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>

**Akarsu 2006** (Continued)

Comparative

Notes First observer results used

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			Low concern
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			Low concern
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		High risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			



**Akarsu 2006** (Continued)

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?	Yes
<b>Could the patient flow have introduced bias?</b>	High risk

**Ashley 1998**
**Study characteristics**

Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars and molars Sealants: excluded Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: UK Setting: extracted teeth Number of participants/teeth/sites: 103 teeth, 103 sites Prevalence: enamel = 0.60, dentine = 0.36
Index tests	Category of test: visual - Downer Sequence of test(s): index tests (ECM, visual, FOTI, radiograph) performed prior to reference standard. Performed in a random order with time delays to ensure independence of tests Examiner training and calibration: 1 examiner Teeth cleaning prior to examination: extracted teeth which had been stored Tooth drying prior to examination: dried for 20 seconds Threshold applied: sound, enamel, dentine
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

**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

**Ashley 1998** (Continued)

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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
Is the reference standards likely to correctly classify the target condition?	Yes		

**Ashley 1998** (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

**Bahrololoomi 2015**

**Study characteristics**

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: no cavitation and early lesions: "intact or had incipient and inconspicuous caries with or without color change were selected"</p> <p>Teeth: primary and permanent molars</p> <p>Sealants: no</p> <p>Restorations: not reported</p> <p>Surface: occlusal</p>
Patient characteristics and setting	<p>Age: 7 to 13 years</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Iran</p> <p>Setting: dental school</p> <p>Number of participants/teeth/sites: 31 participants, 115 teeth (6 of these were excluded "due to patient dropout" so the became 109 teeth)</p> <p>Prevalence: enamel = 0.94, dentine = 0.37</p>
Index tests	<p>Category of test: visual - Ekstrand's visual scoring system</p> <p>Sequence of test(s): index tests (visual, radiograph DIAGNOdent) performed prior to reference standard</p> <p>Examiner training and calibration: not reported</p>

**Bahrololoomi 2015** (Continued)

Teeth cleaning prior to examination: cleaning with a rubber cup and pumice powder

Tooth drying prior to examination: isolation with cotton rolls, and drying

Threshold applied: V0 = no change or slight change of enamel translucency after air drying, V1 = opacity or discolouration distinctly visible after air drying, V2 = opacity or discolouration visible without air drying, V3 = localised enamel breakdown in opaque or discoloured enamel and/or grayish discolouration from the underlying dentine, V4 = cavitation in opaque or discoloured enamel exposing dentine

**Target condition and reference standard(s)**

Category: excavation performed in cases with "obvious or ambiguous dental caries", respectively to assess and score the actual depth of lesions

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: excavated suspicious site

Target condition: sound, enamel, dentine

**Flow and timing**

Participants with index test but no reference standard: unclear whether all surfaces were excavated and if not then what the reference standard was

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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			Low concern
<b>DOMAIN 2: Index Test (All tests)</b>			

**Bahrololoomi 2015** (Continued)

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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		High risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		Low concern
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		High risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		High risk

**Bottenberg 2016**
**Study characteristics**

Patient Sampling	Method of sampling: selected "from a collection extracted for periodontal or orthodontic reasons"
	Included conditions: "a number of sound teeth and a range of occlusal carious lesions" unclear whether this included dentine caries
	Teeth: primary molars
	Sealants: not reported
	Restoration: excluded

**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

**Bottenberg 2016** (Continued)

	Surface: occlusal
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Belgium</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 100 teeth, 169 sites - reported at the tooth level</p> <p>Prevalence: enamel = 0.75, dentine = 0.39</p>
Index tests	<p>Category of test: visual ICDAS</p> <p>Sequence of test(s): visual, then reference standard (photographs also assessed but not used here)</p> <p>Examiner training and calibration: trained and calibrated dentists</p> <p>Teeth cleaning prior to examination: "a batch of extracted cleaned teeth"</p> <p>Tooth drying prior to examination: yes - not details provided</p> <p>Threshold applied: ICDAS criteria</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: following index test, unclear whether same examiners viewed photographs and histology results</p> <p>Training of examiner: unclear</p> <p>Blinding to index test: no</p> <p>Multiple tests: no</p> <p>Site selection: "one to four histological sections and their images were assigned to each examination site"</p> <p>Target condition: caries free, early enamel, deep enamel, outer dentine, dentine, deep dentine</p>
Flow and timing	<p>Participants with index test but no reference standard: 4 - broken during sectioning</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	
<b>Methodological quality</b>	

**Bottenberg 2016** (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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)?			
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			
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		

**Bottenberg 2016** (Continued)

**Could the patient flow have introduced bias?**

Low risk

**Braga 2009**

**Study characteristics**

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: no cavitation and early lesions</p> <p>Teeth: primary molars</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: approximal</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: 84 participants, 131 sites</p> <p>Prevalence: enamel = 0.63, dentine = 0.26</p>
Index tests	<p>Category of test: visual - ICDAS II "the teeth were placed in arch models"</p> <p>Sequence of test(s): index tests (visual, radiograph, laser fluorescence) prior to reference standard</p> <p>Examiner training and calibration: yes, trained</p> <p>Teeth cleaning prior to examination: brush and slurry</p> <p>Tooth drying prior to examination: air dried for 3 seconds</p> <p>Threshold applied: 0: sound tooth, 1: first visual change in enamel, 2: distinct visual change in enamel, 3: localised enamel breakdown, 4: underlying dark shadow from dentine, 5: distinct cavity with visible dentine, 6: extensive distinct cavity with visible dentine</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>



**Braga 2009** (Continued)

Target condition: sound, enamel (outer), enamel (inner), dentine (outer), dentine (inner)

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

Data taken from table 3

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		

**Braga 2009** (Continued)

**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

**Braga 2010**
**Study characteristics**

Patient Sampling	<p>Method of sampling: whole sample was selected in schools or preschools, however those included in DTA part of paper were included due to exfoliation which can be classed as random</p> <p>Included conditions: no cavitation and early lesions; exclusions "teeth with restorations, hypoplastic defects, sealants or frankly cavitated lesions were excluded"</p> <p>Teeth: primary molars</p> <p>Sealants: excluded</p> <p>Restorations: excluded</p> <p>Surface: occlusal</p>
Patient characteristics and setting	<p>Age: 3 to 12 years</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Brazil</p> <p>Setting: "selected from preschools or schools.... and from children who sought dental treatment"</p> <p>Number of participants/teeth/sites: 139 participants, 763 teeth, 50 included in the review because they received a reference standard</p> <p>Prevalence: enamel = 0.52, dentine = 0.26</p>
Index tests	<p>Category of test: 2 visual tests - ICDAS II, Nyvad</p> <p>Sequence of test(s): "performed independent evaluations of all selected occlusal sites, using the NY and ICDAS-II (plus LAA)" randomised and 2 tests separated by 1 week</p>

**Braga 2010** (Continued)

Examiner training and calibration: yes, trained and calibrated

Teeth cleaning prior to examination: brush and slurry

Tooth drying prior to examination: 3-in-1 syringe

Threshold applied:  
ICDAS - 0: sound tooth, 1: first visual change in enamel, 2: distinct visual change in enamel, 3: localised enamel breakdown, 4: underlying dark shadow from dentine, 5: distinct cavity with visible dentine, 6: extensive distinct cavity with visible dentine

Nyvad - 0: sound, 1: active caries (intact surface), 2: active caries (surface discontinuity), 3: active caries (cavity), 4: inactive caries (intact surface), 5: inactive caries (surface discontinuity), 6: inactive caries (cavity)

Target condition and reference standard(s)      Category: histology

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

Training of examiner: "2 examiners blind to each other and to the clinical visual scores"

Blinding to index test: unclear

Multiple tests: no

Site selection: sectioned teeth

Target condition: sound, enamel (outer), enamel (inner), dentine (outer), dentine (inner)

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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		Low risk	

**Braga 2010** (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? 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? 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

**Braun 2017**
**Study characteristics**

Patient Sampling Method of sampling: unclear

**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

**Braun 2017** (Continued)

	<p>Included conditions: "visually intact tooth surfaces"</p> <p>Teeth: permanent molars and premolars</p> <p>Sealants: not reported</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: Germany</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 84 teeth</p> <p>Prevalence: enamel = 0.77, dentine = 0.40</p>
Index tests	<p>Category of test: visual - ICDAS</p> <p>Sequence of test(s): index tests (visual then radiograph) performed prior to reference standard</p> <p>Examiner training and calibration: 2 independent examiners, same examiner for visual and radiograph</p> <p>Teeth cleaning prior to examination: cleaned with a hand scaler</p> <p>Tooth drying prior to examination: not reported</p> <p>Threshold applied: ICDAS codes</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: assumed index test then reference standard but not clearly stated, unlikely that reference standard could have informed index test</p> <p>Training of examiner: experienced examiners</p> <p>Blinding to index test: same examiners as index test</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: "H0" no caries, "H1" caries extending up to outer half of enamel, "H2" caries extending up to inner half of enamel but not into dentine, "H3" caries extending up to outer half of dentine, "H4" caries extending up to inner half of dentine with or without apparent pulp involvement</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>

**Braun 2017** (Continued)

Comparative

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			Low concern
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			High
<b>DOMAIN 4: Flow and Timing</b>			

**Braun 2017** (Continued)

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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Bussanelli 2015**
**Study characteristics**

Patient Sampling	Method of sampling: selected  Included conditions: no cavitation and early lesions "sound or carious primary molars in proximal contact", exclusions "teeth with restoration, occlusal caries, hypoplasias, and an advanced stage of rhizolysis were not included"  Teeth: primary molars  Sealants: not reported  Restorations: excluded  Surface: approximal
Patient characteristics and setting	Age: 5 to 9 years  Sex: not reported  Ethnicity: not reported  Country: Brazil  Setting: paediatric dental clinic  Number of participants/teeth/sites: 45 participants, 59 teeth  Prevalence: enamel = 0.71, dentine = 0.58
Index tests	Category of test: visual - Nyvad  Sequence of test(s): index tests (visual, DIAGNOdent pen, radiograph) prior to reference standard  Examiner training and calibration: experienced  Teeth cleaning prior to examination: brush at low speed, using prophylactic paste and dental floss  Tooth drying prior to examination: unclear  Threshold applied: (0) healthy, (1) active lesion with intact surface, (2) active lesion with discontinuous surface, (3) cavitated active lesion, (4) inactive lesion with intact surface, (5) inactive lesion with discontinuous surface, and (6) cavitated inactive lesion

**Bussanelli 2015** (Continued)

Target condition and reference standard(s)	Category: visual after separation using orthodontic rubber bands (4 mm) for 7 days  Sequence of index test and reference standard: index test then reference standard  Training of examiner: 2 trained and experienced examiners  Blinding to index test: unclear  Multiple tests: no  Site selection: all approximal surfaces  Target condition: healthy, active lesions without loss of structure, signs of caries requiring restoration
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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		



**Bussaneli 2015** (Continued)

<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	Low concern
<b>DOMAIN 3: Reference Standard</b>	
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
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	High risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Bussaneli 2015a**
**Study characteristics**

Patient Sampling	Method of sampling: selected  Included conditions: no cavitation and early lesions "these presented a sound occlusal surface, or had signs suggesting the presence of caries lesion, such as white spots, pigmentations, and cavities"  Teeth: permanent third molars  Sealants: not reported  Restorations: included  Surface: occlusal
Patient characteristics and setting	Age: not reported  Sex: not reported  Ethnicity: not reported  Country: Brazil

**Bussanelli 2015a** (Continued)

Setting: extracted teeth  
 Number of participants/teeth/sites: 77 teeth  
 Prevalence: enamel = 0.63, dentine level not available

Index tests

Category of test: visual - ICDAS  
 Sequence of test(s): index tests (visual, then radiograph) prior to reference standard  
 Examiner training and calibration: experienced examiners "Paediatric Dentistry (Group A), with at least 10 years of experience in clinical practice"  
 Teeth cleaning prior to examination: cleaned with a Robinson brush  
 Tooth drying prior to examination: wet initially, then dried with a triple syringe  
 Threshold applied: ICDAS codes

Target condition and reference standard(s)

Category: histology  
 Sequence of index test and reference standard: index test then reference standard  
 Training of examiner: experienced examiners  
 Blinding to index test: unclear  
 Multiple tests: no  
 Site selection: sectioned teeth  
 Target condition: healthy, outer half of enamel, inner enamel and outer third of dentinal, mid-dentine and inner 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

Focused on group A - experienced examiners, rather than group B - second year dental students; group A greater applicability to our research question

**Methodological quality**

**Item**

**Authors' judgement**

**Risk of bias**

**Applicability concerns**

**DOMAIN 1: Patient Selection**

Was a consecutive or random sample of patients enrolled?

No

**Bussaneli 2015a** (Continued)

Was a case-control design avoided?	Yes	
Did the study avoid inappropriate exclusions?	Yes	
<b>Could the selection of patients have introduced bias?</b>		High risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		High
<b>DOMAIN 2: Index Test (All tests)</b>		
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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		High
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Unclear risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Campos 2017**
**Study characteristics**
**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

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

**Campos 2017** (Continued)

Patient Sampling	<p>Method of sampling: not reported - donated teeth but no indication of consecutive or random sampling</p> <p>Included conditions: no cavitation and unclear on severity of lesions</p> <p>Teeth: permanent: premolars and third molars</p> <p>Sealants: not reported</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: Brazil</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 115 teeth</p> <p>Prevalence: enamel = 0.96, dentine = 0.50</p>
Index tests	<p>Category of test: visual - ICDAS</p> <p>Sequence of test(s): index tests prior to reference standard</p> <p>Examiner training and calibration: "previously trained with the e-learning program provided by the ICDAS Foundation"</p> <p>Teeth cleaning prior to examination: with pumice and rotating brush, washed with distilled water</p> <p>Tooth drying prior to examination: wet initially, then dried with compressed air for 5 seconds</p> <p>Threshold applied: ICDAS codes</p>
Target condition and reference standard(s)	<p>Category: histology - microradiography and stereomicroscopy</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: yes, microradiography used for results</p> <p>Site selection: sectioned teeth</p> <p>Target condition: healthy, outer half of enamel, inner enamel and outer third of dentinal, mid-dentine and inner 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>

**Campos 2017** (Continued)

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Microradiography results reported in analysis

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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)?			
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			

**Campos 2017** (Continued)

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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Carvalho 2018**
**Study characteristics**

Patient Sampling	<p>Method of sampling: selected from a pool of previously extracted teeth</p> <p>Included conditions: unclear "suspected carious lesions on their occlusal surfaces" "those with extensive cavities" were excluded</p> <p>Teeth: permanent third molars</p> <p>Sealants: not reported</p> <p>Restorations: excluded: "sound or restored teeth"</p> <p>Surface: occlusal</p> <p>Final data sample shows sound teeth despite the intention to exclude these</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: 40 teeth</p> <p>Prevalence: enamel = 0.73, dentine = 0.45</p>
Index tests	<p>Category of test: visual - ICDAS</p> <p>Sequence of test(s): index tests (standard visual prior digital images) prior to reference standard</p> <p>Examiner training and calibration: calibration completed, experienced examiners "pediatric dentists with a maximum of 5 years of clinical practice"</p> <p>Teeth cleaning prior to examination: "cleaned with ultrasonic tips to remove calculus and debris"</p> <p>Tooth drying prior to examination: "examined after drying"</p> <p>Threshold applied: ICDAS codes</p>
Target condition and reference standard(s)	<p>Category: microCT: "microCT (1173, Bruker, Kontich, Belgium) with the following acquisition parameters: 70 kV, 114 µA, 7.12- µm pixel size, 1-mm thick Al filter, 1 s exposure, 0.5° rotation step at 360°, and 20-line random movements"</p>

**Carvalho 2018** (Continued)

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

Training of examiner: not reported

Blinding to index test: unclear

Multiple tests: yes, microradiography used for results

Site selection: sectioned teeth

Target condition: healthy, enamel, dentine

"To detect any sign of caries (threshold 1), positive cases were identified by setting a cutoff point of 20% deviation from the mean gray scale level of sound enamel based on the histogram distribution of each specimen. For dentin caries (threshold 2), the cutoff point was set as the demineralized dentin tissue that should be removed from the cavity"

**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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		

**Carvalho 2018** (Continued)

<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	Low concern
<b>DOMAIN 3: Reference Standard</b>	
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
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Castilho 2016**
**Study characteristics**

Patient Sampling	Method of sampling: consecutive Included conditions: no cavitation and early lesions Teeth: third molars, requiring extraction Sealants: no Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: 16 to 39 years Sex: 10 male, 16 female Ethnicity: not reported Country: Brazil



**Castilho 2016** (Continued)

	Setting: dental clinic
	Number of participants/teeth/sites: 26 participants, 43 teeth
	Prevalence: enamel = 0.81, dentine = 0.07
Index tests	Category of test: visual - ICDAS  Sequence of test(s): index tests (visual then DIAGNOdent pen) prior to reference standard  Examiner training and calibration: yes  Teeth cleaning prior to examination: brush and slurry  Tooth drying prior to examination: yes  Threshold applied: ICDAS - 0: sound tooth, 1: first visual change in enamel, 2: distinct visual change in enamel, 3: localised enamel breakdown, 4: underlying dark shadow from dentine, 5: distinct cavity with visible dentine, 6: extensive distinct cavity with visible dentine
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  Target condition: healthy, 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	
<b>Methodological quality</b>	
<b>Item</b>	<b>Authors' judgement</b>
	<b>Risk of bias</b>
	<b>Applicability concerns</b>
<b>DOMAIN 1: Patient Selection</b>	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes

**Castilho 2016** (Continued)

<b>Could the selection of patients have introduced bias?</b>	Low risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>	Low concern
<b>DOMAIN 2: Index Test (All tests)</b>	
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
<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	Low concern
<b>DOMAIN 3: Reference Standard</b>	
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
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Chawla 2012**
**Study characteristics**

Patient Sampling	Method of sampling: selected
	Included conditions: no cavitation and early lesions
	Teeth: primary molars

**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

**Chawla 2012** (Continued)

	<p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Australia</p> <p>Setting: extracted teeth - "mesial surfaces of first primary molars were placed in contact with distal surfaces of second primary molars"</p> <p>Number of participants/teeth/sites: 135 sites</p> <p>Prevalence: enamel = 0.61, dentine = 0.24</p>
Index tests	<p>Category of test: visual - modified ICDAS II, teeth set in a model to recreate approximal contact</p> <p>Sequence of test(s): index tests (visual, radiograph, then DIAGNOdent then DIAGNOdent pen) prior to reference standard</p> <p>Examiner training and calibration: training completed</p> <p>Teeth cleaning prior to examination: brush and slurry</p> <p>Tooth drying prior to examination: yes</p> <p>Threshold applied: ICDAS II codes 0–2 remained as codes 0–2, codes 3 and 4 were collapsed as code 3, and codes 5 and 6 were collapsed as code 4</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: healthy, inner/outer enamel, inner/outer 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	Data used from examiner 1

**Chawla 2012** (Continued)

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			
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		

**Chawla 2012** (Continued)

**Could the patient flow have introduced bias?**

Low risk

**Cinar 2013**

**Study characteristics**

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: no cavitation and early lesions</p> <p>Teeth: primary molars - pre and post-exfoliation</p> <p>Sealants: excluded</p> <p>Restorations: excluded</p> <p>Surface: occlusal</p>
Patient characteristics and setting	<p>Age: 9 to 11 years</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Turkey</p> <p>Setting: dental hospital</p> <p>Number of participants/teeth/sites: 44 sites</p> <p>Prevalence: enamel = 0.75, dentine = 0.20</p>
Index tests	<p>Category of test: visual - no criteria specified</p> <p>Sequence of test(s): index tests (visual, radiograph then DIAGNOdent and DIAGNOdent pen) prior to reference standard</p> <p>Examiner training and calibration: not reported</p> <p>Teeth cleaning prior to examination: cleaned with paste</p> <p>Tooth drying prior to examination: air spray was available</p> <p>Threshold applied: 0 = no or slight change in enamel translucency after prolonged air (&gt; 5 seconds), 1 = opacity or white and brown discolouration hardly visible on the wet surface, but distinctly visible after air drying, 2 = opacity or white and brown discolouration distinctly visible without air drying, 3 = localised enamel breakdown in opaque or discoloured enamel and/or grayish discolouration from the underlying dentine, 4 = cavitation in opaque or discoloured enamel exposing the dentine</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>

**Cinar 2013** (Continued)

Site selection: sectioned teeth

Target condition: sound, outer enamel, inner 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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			Low concern
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			Low concern
<b>DOMAIN 3: Reference Standard</b>			
Is the reference standards likely to correctly classify the target condition?	Yes		



**Costa 2002** (Continued)

	Tooth drying prior to examination: not reported  Threshold applied: no caries, caries confined to enamel, caries extending into dentine
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: 1 (damaged during sectioning)  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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		



**Costa 2002** (Continued)

<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	High
<b>DOMAIN 3: Reference Standard</b>	
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
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Da Silva 2010**

<b>Study characteristics</b>	
Patient Sampling	Method of sampling: selected Included conditions: early/enamel lesions, no restored teeth Teeth: primary first and second upper and lower molars Sealants: not reported Restorations: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 50 teeth, 50 sites

**Da Silva 2010** (Continued)

Prevalence: enamel = 0.56, dentine = 0.38

Index tests

Category of test: visual - Ekstrand

Sequence of test(s): 1) radiography, 2) visual, 3) histology

Examiner training and calibration: 2 dental radiologists with minimum of 5 years experience. To ensure calibration, a training exercise was completed before each observation session. Instructions were provided and observers became familiar with the image scoring to be used

Teeth cleaning prior to examination: extracted teeth

Tooth drying prior to examination: performed using air/water spray

Threshold applied: Ekstrand criteria

Target condition and reference standard(s)

Category: histology

Sequence of index test and reference standard: reference standard last

Training of examiner: done by an oral pathologist

Blinding to index test: yes - different examiner (histology assumed to be blind regardless)

Multiple tests: no

Site selection: teeth were embedded into acrylic resin and sectioned in a mesiodistal direction

Target condition: D0 = no caries, D1 = caries lesion limited to the outer half of the enamel, D2 = caries extending into inner half of the enamel, but not to amelodentinal junction, D3 = caries limited to the outer half of the dentine, D4 = caries involving the inner half of the 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

Data calculated from sensitivity/specificity and prevalence reported in paper; we used the data for observer 1

**Methodological quality**

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

**Da Silva 2010** (Continued)

<b>Could the selection of patients have introduced bias?</b>		High risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		High
<b>DOMAIN 2: Index Test (All tests)</b>		
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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		High
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Diniz 2009**
**Study characteristics**

Patient Sampling	Method of sampling: selected
------------------	------------------------------

**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

**Diniz 2009** (Continued)

	<p>Included conditions: unclear severity of disease</p> <p>Teeth: permanent 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: Switzerland</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 163 teeth/sites</p> <p>Prevalence: enamel = 0.96, dentine = 0.47</p>
Index tests	<p>Category of test: visual - ICDAS II</p> <p>Sequence of test(s): index prior to reference standard</p> <p>Examiner training and calibration: 2 experienced dentists with training</p> <p>Teeth cleaning prior to examination: yes</p> <p>Tooth drying prior to examination: moist then dried</p> <p>Threshold applied: ICDAS II</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, 2 senior researchers with consensus</p> <p>Blinding to index test: yes</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: sound, inner/outer enamel, inner/outer 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	

**Diniz 2009** (Continued)

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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)?			
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			
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		

**Diniz 2009** (Continued)

**Could the patient flow have introduced bias?**

Low risk

**Diniz 2011**

**Study characteristics**

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: sound, early/enamel lesions, cavitated (level of cavitation not reported), no sealants, no restorations, no hypoplasia, no fluorosis, no caries lesions on smooth or proximal surfaces</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: Brazil</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 104 teeth, 104 sites</p> <p>Prevalence: enamel = 0.94, dentine = not possible to calculate due to no clear threshold between enamel and dentine caries</p>
Index tests	<p>Category of test: visual - ICDAS</p> <p>Sequence of test(s): 1) visual, 2) radiography, 3) treatment decision (visual and radiography scores made available to the examiner), 4) histology. Although same examiners also did visual tests: "The examiners were blinded during the x-ray examination, and they did not have access to the test tooth"</p> <p>Examiner training and calibration: 4 dentists from Department of Paediatric Dentistry: 2 adjunct professors with 15 years' postgraduate experience and 2 PhD students with 8 years' postgraduate experience. Training carried out for ICDAS but not for radiography</p> <p>Teeth cleaning prior to examination: "Calculus and debris were removed using a scaler, and the teeth were then cleaned for 15 seconds"</p> <p>Tooth drying prior to examination: 3-in-1 air syringe</p> <p>Threshold applied: ICDAS codes</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: 2 experienced examiners and any disagreements re-examined and consensus reached</p> <p>Blinding to index test: yes - assumed as it is histology</p>

**Diniz 2011** (Continued)

Multiple tests: no

Site selection: teeth longitudinally sectioned through centre of each test site resulting in 2 sections corresponding to the test site. The section with a more severe lesion was analysed

Target condition: (D0) no caries, (D1) caries lesion limited to outer half of enamel, (D2) caries extending into inner half of enamel or outer half of dentine, (D3) caries limited to middle third of dentine, and (D4) caries involving inner half of 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

Data calculated from sensitivity/specificity (mean of the 4 assessors was reported) and prevalence reported in paper

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	

**Diniz 2011** (Continued)

**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

**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

**Diniz 2012**
**Study characteristics**

Patient Sampling

Method of sampling: selected: "88 patients who each had at least one posterior tooth scheduled for extraction"

Included conditions: no cavitation and early lesions "ranged from having macroscopically intact occlusal surfaces to different stages of noncavitated and cavitated carious lesions"

Teeth: permanent molars and premolars

Sealants: excluded

Restorations: excluded

Surface: occlusal

Patient characteristics and setting

Age: 18 to 35 years

Sex: not reported

Ethnicity: not reported

**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

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



**Diniz 2012** (Continued)

Country: Brazil

Setting: index test performed in clinical setting at dental school, then teeth extracted for reference standard

Number of participants/teeth/sites: 88 participants, 105 teeth

Prevalence: enamel = 0.95, dentine = 0.26

## Index tests

Category of test: visual - ICDAS

Sequence of test(s): index tests (visual, radiograph, DIAGNOdent, DIAGN-Odent pen, and VistaProof) followed by reference standard

Examiner training and calibration: 1 experienced examiner - ICDAS training and calibration

Teeth cleaning prior to examination: low-speed handpiece with a rotating brush and water

Tooth drying prior to examination: unclear

Threshold applied: ICDAS thresholds, only ICDAS 0-4 found in sample

## Target condition and reference standard(s)

Category: histology

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

Training of examiner: 2 trained examiners

Blinding to index test: unclear

Multiple tests: no

Site selection: marked on photographs then sectioned teeth

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?

Unclear

Was a case-control design avoided?

Yes

**Diniz 2012** (Continued)

Did the study avoid inappropriate exclusions?	Yes	
<b>Could the selection of patients have introduced bias?</b>		Unclear risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		Low concern
<b>DOMAIN 2: Index Test (All tests)</b>		
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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		Low concern
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Diniz 2019**
**Study characteristics**

Patient Sampling	Method of sampling: unclear
------------------	-----------------------------

**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

**Diniz 2019** (Continued)

	<p>Included conditions: unclear on severity of lesions "with varying conditions from sound to that of different stages of carious lesion"</p> <p>Teeth: primary 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: Brazil</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 88 teeth</p> <p>Prevalence: enamel = 0.75, dentine = 0.63</p>
Index tests	<p>Category of test: visual - ICDAS</p> <p>Sequence of test(s): index tests (visual appears to be conducted prior to: DIAGNOdent, DIAGNOdent pen, QLF, and MidWest) followed by reference standard</p> <p>Examiner training and calibration: 1 trained and experienced examiner</p> <p>Teeth cleaning prior to examination: rinsed with 3-in-1 syringe</p> <p>Tooth drying prior to examination: "the teeth were assessed moist and subsequently after drying for 5 s"</p> <p>Threshold applied: ICDAS codes</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, 2 senior researchers with consensus</p> <p>Blinding to index test: yes, independent assessors</p> <p>Multiple tests: no</p> <p>Site selection: marked on photographs then sectioned teeth</p> <p>Target condition: sound, inner/outer enamel, inner/outer 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>

**Diniz 2019** (Continued)

Comparative

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			
Was there an appropriate interval between index test and reference standard?	Yes		

**Diniz 2019** (Continued)

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

**Ekstrand 2007**
**Study characteristics**

Patient Sampling	<p>Method of sampling: teeth were collected from clinics and the lesions selected</p> <p>Included conditions: not reported, suggestion that sound sites were not selected</p> <p>Teeth: primary teeth</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: occlusal and smooth/approximal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Scotland and Denmark</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 141 teeth</p> <p>Prevalence: enamel = 0.72, dentine = 0.38</p>
Index tests	<p>Category of test: visual - ICDAS II</p> <p>Sequence of test(s): index tests followed by reference standard, only reporting on study 1 for this review</p> <p>Examiner training and calibration: not reported, but known to be experienced ICDAS examiners</p> <p>Teeth cleaning prior to examination: yes</p> <p>Tooth drying prior to examination: not reported</p> <p>Threshold applied: ICDAS II codes</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>

**Ekstrand 2007** (Continued)

Multiple tests: no  
 Site selection: sectioned teeth  
 Target condition: sound, inner/outer enamel, inner/outer dentine

## Flow and timing

Participants with index test but no reference standard: 2 - broken during sectioning, no risk of bias  
 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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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)?			
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
Is the reference standards likely to correctly classify the target condition?	Yes		

**Ekstrand 2007** (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Ekstrand 2011**
**Study characteristics**

Patient Sampling	<p>Method of sampling: selected, "collected from plastic boxes of thousands of extracted and exfoliated teeth"</p> <p>Included conditions: no cavitation and early lesions, "ranged from having macroscopically intact occlusal surfaces to different stages of noncavitated and cavitated carious lesions"</p> <p>Teeth: primary: molars, canines and incisors, and permanent: molars, premolars and canines</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Colombia</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: "132 sound/carious proximal surfaces from 106 primary teeth and 160 sound/carious proximal surfaces from 140 permanent teeth"</p> <p>Prevalence:</p> <ul style="list-style-type: none"> <li>- primary: enamel = 0.82, dentine = 0.64</li> <li>- permanent: enamel = 0.78, dentine = 0.33</li> </ul>

**Ekstrand 2011** (Continued)

Index tests	Category of test: visual - ICDAS Teeth not set up in a model to recreate approximal contact Sequence of test(s): index tests (visual, radiograph) prior to reference standard (the order of teeth and radiographs in the 1st examination was changed in the 2nd) Examiner training and calibration: 3 examiners trained on a 3-day ICDAS course Teeth cleaning prior to examination: yes Tooth drying prior to examination: 3-in-1 syringe used Threshold applied: ICDAS codes
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: experienced examiner Blinding to index test: unclear Multiple tests: no Site selection: marked on photographs then sectioned teeth Target condition: 0 = sound surfaces, 1 = demineralization in the outer half of the enamel, 2 = demineralization involving the area between the inner half of the enamel and outer third of the dentine, 3 = demineralization reaching the middle third of the dentine, 4 = demineralization involving the inner third of the dentine
Flow and timing	Participants with index test but no reference standard: "Due to breakdown of the tooth under the preparation, the histological examination was done on 112 primary surfaces and on 151 permanent surfaces" 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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	



**Ekstrand 2011** (Continued)

<b>Are there concerns that the included patients and setting do not match the review question?</b>	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
--	----

<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Low risk
--	----------

<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	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
--	---------

<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
---	----------

<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	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
---	----

<b>Could the patient flow have introduced bias?</b>	Low risk
---	----------

**Ekstrand 2011a**
**Study characteristics**

**Ekstrand 2011a** (Continued)

Patient Sampling	Same as previous study, entered to allow for data entry of permanent teeth
Patient characteristics and setting	
Index tests	
Target condition and reference standard(s)	
Flow and timing	
Comparative	
Notes	

**Erten 2005**

**Study characteristics**

Patient Sampling	<p>Method of sampling: "selected following examination under a stereomicroscope"</p> <p>Included conditions: no cavitation and early lesions, "no signs of demineralization to those with varying degrees of demineralization"</p> <p>Teeth: permanent 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: Turkey</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 84 teeth</p> <p>Prevalence: enamel = 0.63, dentine = 0.23</p>
Index tests	<p>Category of test: visual - ERK</p> <p>Sequence of test(s): index tests (visual, intraoral camera, (Rydalmer NSW 2116, Australia) and operating microscope (Moller-Wedel, Dento 300, Germany)) prior to reference standard, all index tests separated by at least 1 week</p> <p>Examiner training and calibration: 4 examiners all with at least 4 years' experience</p> <p>Teeth cleaning prior to examination: cleaned with a toothbrush</p> <p>Tooth drying prior to examination: air and water from the unit air-water syringe</p>

**Erten 2005** (Continued)

Threshold applied: ERK criteria

Target condition and reference standard(s)	Category: histology - stereomicroscope  Sequence of index test and reference standard: index test then reference standard  Training of examiner: 2 experienced examiners - possibly same examiners as index test  Blinding to index test: unclear  Multiple tests: no  Site selection: sectioned teeth  Target condition: 0 = sound surfaces, 1 = demineralization in the outer half of the enamel, 2 = demineralization involving the area between the inner half of the enamel and outer third of the dentine, 3 = demineralization reaching the middle third of the dentine, 4 = demineralization involving the inner third of the 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	Mean value of 4 examiners results used in data extraction

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		

**Erten 2005** (Continued)

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

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

**Forgie 2003**
**Study characteristics**

Patient Sampling	Method of sampling: selected - methods not reported Included conditions: not reported Teeth: permanent premolars and molars Sealants: not reported Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported

**Forgie 2003** (Continued)

	Ethnicity: not reported Country: UK Setting: extracted teeth Number of participants/teeth/sites: 40 molars, 40 premolars, (20 canines not assessed) Prevalence: enamel = 0.86, dentine = 0.49		
Index tests	Category of test: visual - other Sequence of test(s): index tests (visual, intraoral camera (Clear-Vu, UK)) prior to reference standard, all index tests separated by at least 1 week Examiner training and calibration: 6 examiners not calibrated Teeth cleaning prior to examination: not reported Tooth drying prior to examination: not reported - examiners followed their usual clinical practice Threshold applied: sound, enamel lesion, enamel cavity, dentine lesion, dentine cavity		
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: 2 trained assessors Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: same system as index test		
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	Mean results used for analysis		
<b>Methodological quality</b>			
<b>Item</b>	<b>Authors' judgement</b>	<b>Risk of bias</b>	<b>Applicability concerns</b>
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		

**Forge 2003** (Continued)

Was a case-control design avoided?	Yes	
Did the study avoid inappropriate exclusions?	Unclear	
<b>Could the selection of patients have introduced bias?</b>		High risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		High
<b>DOMAIN 2: Index Test (All tests)</b>		
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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		High
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Freitas 2016**
**Study characteristics**

Patient Sampling	Method of sampling: selected - convenience sample in basic health unit
------------------	--

**Freitas 2016** (Continued)

	<p>Included conditions: "The approximal surfaces ranged from sound to different stages of noncavitated and cavitated carious lesions"</p> <p>Teeth: "primary molar that was in an advanced stage of root resorption" (pre and post-exfoliation)</p> <p>Sealants: not reported</p> <p>Restorations: excluded</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: 8 to 12 years</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Brazil</p> <p>Setting: health unit</p> <p>Number of participants/teeth/sites: 56 participants, 89 teeth, 166 surfaces</p> <p>Prevalence: enamel = 0.68, dentine - not possible to calculate as threshold combines inner enamel and outer third of dentine, highest prevalence would be 0.5</p>
Index tests	<p>Category of test: visual - ICDAS - before and after tooth separation</p> <p>Sequence of test(s): index tests (visual then radiograph) on 2 occasions but the same BW was used from the patient's records; performed prior to reference standard</p> <p>Examiner training and calibration: experienced researcher, who was previously trained and calibrated, same examiner for visual and radiograph</p> <p>Teeth cleaning prior to examination: "performed after professional prophylaxis using a low-speed handpiece with a rotating bristle brush and pumice/water slurry. The approximal surfaces were also flossed for better surface evaluation"</p> <p>Tooth drying prior to examination: not reported</p> <p>Threshold applied: ICDAS codes</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: 2 experienced and trained examiners, who did not participate in the clinical examinations, evaluated each tooth section independently</p> <p>Blinding to index test: separate examiner to index test</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: 0 = no caries, 1 = carious lesion limited to the outer half of the enamel, 2 = caries extending into the inner half of the enamel or outer half of the dentine, 3 = caries limited to the middle third of the dentine, and 4 = caries involving the inner half of the 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>

**Freitas 2016** (Continued)

Time interval between tests: "monitored until natural exfoliation two weeks later; or scheduled for extraction" therefore unclear on time delay between index tests and reference standard

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Concern regarding unclear threshold between enamel and dentine lesions, since we are reporting at the enamel level the applicability remains low concern for the reference standard. Data extracted for the index test prior to separation as more clinically relevant

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			Low concern
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			Low concern
<b>DOMAIN 3: Reference Standard</b>			
Is the reference standards likely to correctly classify the target condition?	Yes		



**Freitas 2016** (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?** 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? Unclear

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

**Could the patient flow have introduced bias?** Unclear risk

**Goel 2009**
**Study characteristics**

Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early enamel lesions Teeth: first and second molars Sealants: excluded Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: 8 to 12 years Sex: not reported Ethnicity: not reported Country: India Setting: index test performed in a clinical setting prior to extraction Number of participants/teeth/sites: 84 teeth, 83 sites Prevalence: enamel = 0.54, dentine = 0.43
Index tests	Category of test: visual - Ekstrand Sequence of test(s): index tests (radiograph, visual, and DIAGNOdent) performed prior to reference standard

**Goel 2009** (Continued)

Examiner training and calibration: not reported

Teeth cleaning prior to examination: calculus removed with scaler

Tooth drying prior to examination: air dried 5 seconds

Threshold applied: presence or absence of carious lesions was recorded using Ekstrand's criteria

Note - radiograph results may have influenced visual assessment

## 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, inner/outer enamel, inner/outer dentine

## Flow and timing

Participants with index test but no reference standard: 1

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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			Low concern
<b>DOMAIN 2: Index Test (All tests)</b>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

**Goel 2009** (Continued)

If a threshold was used, was it pre-specified?	Yes
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear
<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Unclear risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	Low concern
<b>DOMAIN 3: Reference Standard</b>	
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
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Haak 2002**
**Study characteristics**

Patient Sampling	Method of sampling: unclear, "obtained for research purposes" Included conditions: no cavitation and early enamel lesions Teeth: permanent premolars, first and second molars Sealants: not reported Restorations: excluded Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported

**Haak 2002** (Continued)

	Country: Germany
	Setting: extracted teeth
	Number of participants/teeth/sites: 160 teeth, 320 sites
	Prevalence: enamel = 0.67, dentine = 0.37
Index tests	<p>Category of test: visual - other - viewed outside of a model set-up</p> <p>Sequence of test(s): index tests (visual, then magnifying aid and prism loupe) performed prior to reference standard - order was randomised and assessments separated by 2 weeks, however the part of the study used in the review focusses on the 2 examiners who viewed the teeth outside of the jaw model</p> <p>Examiner training and calibration: 2 examiners with consensus, no training described</p> <p>Teeth cleaning prior to examination: not reported</p> <p>Tooth drying prior to examination: not reported</p> <p>Threshold applied: "0 = no lesion, 1 = enamel opacity with smooth surface, 2 = enamel opacity with rough surface, 3 = cavitation restricted to the enamel, 4 = cavitation extending into dentine"; "instructed to assess only lesions coronal to the cemento-enamel junction"</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: same assessors as index test</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: (a) presence of caries; (b) presence of cavitation</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	
<b>Methodological quality</b>	
<b>Item</b>	<b>Authors' judgement</b>
	<b>Risk of bias</b>
	<b>Applicability concerns</b>
<b>DOMAIN 1: Patient Selection</b>	

**Haak 2002** (Continued)

Was a consecutive or random sample of patients enrolled?	Unclear	
Was a case-control design avoided?	Yes	
Did the study avoid inappropriate exclusions?	Unclear	
<b>Could the selection of patients have introduced bias?</b>		Unclear risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		High
<b>DOMAIN 2: Index Test (All tests)</b>		
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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		High
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Hintze 2003**
**Study characteristics**

Patient Sampling	Method of sampling: selected Included conditions: not reported Teeth: canines, premolars and molars Sealants: not reported Restorations: excluded Surface: approximal and occlusal surfaces in each tooth were scored
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Denmark Setting: extracted teeth Number of participants/teeth/sites: 198 teeth: 373 approximal sites and 158 occlusal Prevalence: enamel = 0.26, dentine = 0.13
Index tests	Category of test: visual - scale not specified, no model set up to recreate approximal contact of teeth Sequence of test(s): visual, radiograph then histology Examiner training and calibration: not reported Teeth cleaning prior to examination: "cleaned with rotating instruments and pumice stone" Tooth drying prior to examination: "teeth were dried and assessed by a dentist" Threshold applied: sound, enamel, outer/inner dentine
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: reference standard followed index tests Training of examiner: not reported Blinding to index test: yes - assumed as it is histology Multiple tests: no Site selection: serially sectioned in 700 µm thick planes Target condition: sound, caries in enamel, and caries in 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

**Hintze 2003** (Continued)

Participants receiving both tests but excluded from results: 0

Comparative

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			

**Hintze 2003** (Continued)

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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Hintze 2003a**
**Study characteristics**

Patient Sampling	Occlusal surfaces
Patient characteristics and setting	
Index tests	
Target condition and reference standard(s)	
Flow and timing	
Comparative	
Notes	

**Huth 2010**
**Study characteristics**

Patient Sampling	Method of sampling: randomised Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: excluded Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: dental hospital Number of participants/teeth/sites: 117 participants



**Huth 2010** (Continued)

	Prevalence: enamel = 0.66, dentine = 0.37
Index tests	Category of test: visual - Ekstrand  Sequence of test(s): index tests (visual, radiograph, DIAGNOdent pen) performed prior to reference standard  Examiner training and calibration: unclear, completed by a dental professional  Teeth cleaning prior to examination: yes  Tooth drying prior to examination: yes  Threshold applied: Ekstrand criteria
Target condition and reference standard(s)	Category: excavation where suspected of cavitation, otherwise radiographs and visual were combined  Sequence of index test and reference standard: index test then reference standard  Training of examiner: experienced, same examiners as index tests or aware of the results of the index test  Blinding to index test: unclear  Multiple tests: yes  Site selection: unclear which site was investigated with which test  Target condition: sound, enamel, dentinal lesions
Flow and timing	Participants with index test but no reference standard: 0  Participants with reference standard but no index test: 21  Time interval between tests: minimal  Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data used from table 2, n/a results excluded from data

**Methodological quality**

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

**Huth 2010** (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? 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?** 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

**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

**Iranzo-Cortes 2017**
**Study characteristics**

Patient Sampling	Method of sampling: selected Included conditions: unclear Teeth: permanent premolars and molars Sealants: unclear
------------------	--

**Iranzo-Cortes 2017** (Continued)

	Restorations: unclear
	Surface: occlusal
Patient characteristics and setting	Age: 18 to 55 years Sex: not reported Ethnicity: not reported Country: Spain Setting: extracted teeth Number of participants/teeth/sites: 65 teeth Prevalence: enamel = 0.77, dentine = 0.17
Index tests	Category of test: visual - ICDAS II Sequence of test(s): index tests performed (visual then DIAGN-Odent) prior to reference standard Examiner training and calibration: 35 teeth used for calibration Teeth cleaning prior to examination: calculus and residues were removed from the selected teeth, using a KAVO Sonic Flex Tooth drying prior to examination: triple air syringe was used to dry teeth Threshold applied: ICDAS codes
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: ss marked prior to index test, then sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 1 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	
<b>Methodological quality</b>	
<b>Item</b>	<b>Authors' judgement</b>
	<b>Risk of bias</b>
	<b>Applicability concerns</b>

**Iranzo-Cortes 2017** (Continued)

**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	
<b>Could the selection of patients have introduced bias?</b>		High risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Iranzo-Cortes 2018**

**Study characteristics**

Patient Sampling	<p>Method of sampling: selected - "teeth extracted for orthodontic or periodontal reasons was selected"</p> <p>Included conditions: "healthy or present incipient caries lesions but those with large cavitated lesions or filled surfaces were excluded"</p> <p>Teeth: permanent premolars and molars</p> <p>Sealants: unclear</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: Spain</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 65 teeth</p> <p>Prevalence: enamel = 0.68, dentine not reported</p>
Index tests	<p>Category of test: visual - ICDAS II</p> <p>Sequence of test(s): index tests performed (visual then VistaProof) prior to reference standard</p> <p>Examiner training and calibration: 35 teeth used for calibration</p> <p>Teeth cleaning prior to examination: calculus and residues were removed from the selected teeth, using a KAVO Sonic Flex</p> <p>Tooth drying prior to examination: triple air syringe was used to dry teeth</p> <p>Threshold applied: ICDAS codes</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: marked prior to index test, then sectioned teeth</p> <p>Target condition: sound, inner/outer enamel, inner/outer 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>



**Iranzo-Cortes 2018** (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?	Yes
<b>Could the patient flow have introduced bias?</b>	Low risk

**Jablonski-Momeni 2008**
**Study characteristics**

Patient Sampling	<p>Method of sampling: selected, "One to four easily re-located sites within the pit and fissure system of each tooth were chosen as potential investigation sites"</p> <p>Included conditions: not reported</p> <p>Teeth: permanent premolars and molars</p> <p>Sealants: unclear</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: Germany</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 100 teeth, 181 potential sites</p> <p>Prevalence: enamel = 0.76, dentine = 0.46</p>
Index tests	<p>Category of test: visual - ICDAS</p> <p>Sequence of test(s): index tests performed prior to reference standard</p> <p>Examiner training and calibration: training provided by experienced examiner and calibration performed on 20 teeth</p> <p>Teeth cleaning prior to examination: cleaned carefully with a rotating brush and water</p> <p>Tooth drying prior to examination: dried with a 3-in-1 syringe</p> <p>Threshold applied: ICDAS codes</p>
Target condition and reference standard(s)	Category: histology

**Jablonski-Momeni 2008** (Continued)

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

Training of examiner: 4 experienced examiners

Blinding to index test: same examiners as index test

Multiple tests: 2 scales used, Downer and ERK

Site selection: number of sections taken dependent on the size of lesion

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

## Flow and timing

Participants with index test but no reference standard: 88 missing due to "section damage on some teeth and to some not being scored by all examiners"

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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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)?			
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	



**Jablonski-Momeni 2008** (Continued)

**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

**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

**Jablonski-Momeni 2012**
**Study characteristics**

Patient Sampling	Method of sampling: unclear Included conditions: no cavitation and early lesions Teeth: permanent premolars and molars Sealants: not reported Restorations: excluded Surface: occlusal, "permanent posterior teeth without occlusal restorations"
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Germany Setting: extracted teeth Number of participants/teeth/sites: 84 teeth

**Jablonski-Momeni 2012** (Continued)

	Prevalence: enamel = 0.84, dentine = 0.48
Index tests	Category of test: visual - ICDAS Sequence of test(s): index tests (visual then DIAGNOdent and Vista-Cam) performed prior to reference standard Examiner training and calibration: 2 examiners, "doctoral student calibrated by an experienced investigator" Teeth cleaning prior to examination: yes, method not reported Tooth drying prior to examination: not reported Threshold applied: ICDAS codes
Target condition and reference standard(s)	Category: excavation - of all teeth 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: all teeth opened with rotating instrument 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: 4, "While using the FC device, 4 investigation sites could not be assessed due to technical problems" 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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		Unclear risk	

**Jablonski-Momeni 2012** (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? 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?** 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

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? No

**Could the patient flow have introduced bias?** Low risk

**Kim 2017**
**Study characteristics**

Patient Sampling	Method of sampling: not clearly reported
	Included conditions: severity of condition unclear: "subjects with 1 or more proximal caries surfaces detected visually or radiographically were included in the study", restorations were included
	Teeth: permanent molars and premolars

**Kim 2017** (Continued)

	<p>Sealants: not reported</p> <p>Restorations: included</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: 19 to 60 years</p> <p>Sex: 55% male</p> <p>Ethnicity: not reported</p> <p>Country: South Korea</p> <p>Setting: clinical setting</p> <p>Number of participants/teeth/sites: 65 participants, 280 sites</p> <p>Prevalence: enamel = 0.61, dentine = 0.20</p>
Index tests	<p>Category of test: visual - modified ICDAS</p> <p>Sequence of test(s): visual then radiograph followed by QLF, radiograph was the reference standard</p> <p>Examiner training and calibration: 1 trained examiner completed all index tests and reference standard</p> <p>Teeth cleaning prior to examination: full-mouth scaling and polishing</p> <p>Tooth drying prior to examination: not reported</p> <p>Threshold applied: "V0: sound, V1: white or brown spot, V2: localised enamel breakdown, V3: underlying dark shadow or colour change, and V4: clinically cavitated or exposed dentine"</p>
Target condition and reference standard(s)	<p>Category: radiograph</p> <p>Sequence of index test and reference standard: reference standard prior to index test</p> <p>Training of examiner: not reported, but experienced</p> <p>Blinding to index test: no</p> <p>Multiple tests: no</p> <p>Site selection: approximal surfaces</p> <p>Target condition: sound, outer/inner enamel, outer/inner dentine</p>
Flow and timing	<p>Participants with index test but no reference standard: 15</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>Data used for the fluorescence images method as the 2 x 2 figures were not available for the software method</p>

Kim 2017 (Continued)

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			Low concern
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		High risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)

131

**Kim 2017** (Continued)

Were all patients included in the analysis? No

**Could the patient flow have introduced bias?** High risk

**Ko 2015**

**Study characteristics**

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: severity of condition unclear: "proximal surfaces with extensive cavities involving more than half of the proximal surface were excluded"</p> <p>Teeth: permanent molars and premolars</p> <p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: South Korea</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 100 teeth (5 were damaged so only 95 reported in results)</p> <p>Prevalence: enamel = 0.80, dentine = 0.15</p>
Index tests	<p>Category of test: visual - ICDAS (teeth set in pairs to model approximal contact)</p> <p>Sequence of test(s): visual then radiograph followed by QLF</p> <p>Examiner training and calibration: 1 calibrated dentist</p> <p>Teeth cleaning prior to examination: cleaned of all soft tissues</p> <p>Tooth drying prior to examination: dried with cotton wool</p> <p>Threshold applied: ICDAS codes</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>

**Ko 2015** (Continued)

Site selection: sectioned teeth

Target condition: "enamel demineralization or a narrow surface zone of opacity", enamel or outer/inner dentine caries

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

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			Low concern
<b>DOMAIN 3: Reference Standard</b>			
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		

**Ko 2015** (Continued)

**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

**Kockanat 2017**
**Study characteristics**

Patient Sampling	Method of sampling: unclear  Included conditions: non-cavitated; "occlusal surfaces of the teeth had minimal macroscopic destruction"  Teeth: primary molars  Sealants: unclear  Restorations: excluded  Surface: occlusal
Patient characteristics and setting	Age: 9 to 12 years  Sex: not reported  Ethnicity: not reported  Country: Turkey  Setting: in vivo study conducted in dental hospital, followed by in vitro after extraction  Number of participants/teeth/sites: 120 teeth (144 teeth were examined and measurements made with caries detection devices, but 120 of the 144 teeth were reported; due to inconsistencies in caries measurement results), clarification provided by study author  Prevalence: enamel = 0.78, dentine = 0.32
Index tests	Category of test: visual - ICDAS  Sequence of test(s): visual, Sopro, DIAGNOdent pen then CarieScan PRO  Examiner training and calibration: unclear, 2 independent examiners  Teeth cleaning prior to examination: plaque removed, washed without pumice



**Kockanat 2017** (Continued)

Tooth drying prior to examination: air water spray, dried again for 5 seconds prior to DD

Threshold applied:

- DIAGNOdent pen: 0 to 13 = sound, 14 to 29 = enamel, 30+ = dentine

- SoproCam: (0) no visible radiolucency, (1) radiolucency in the enamel, (2) radiolucency in the dentine, involving the surface or the outer third of the dentine, and (3) radiolucency in the dentine, involving the inner third of the dentine

Device specifics: cylinder sapphire tip for DIAGNOdent pen, "The images were recorded to Sopro Imaging program and evaluated according to the criteria of Rechmann"

Target condition and reference standard(s)	Category: histology  Sequence of index test and reference standard: index tests then reference standard  Training of examiner: not reported  Blinding to index test: not reported  Multiple tests: no  Site selection: sectioned teeth  Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep dentine
Flow and timing	Participants with index test but no reference standard: 24  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 for examiner 1, with the comparison of in vivo index test versus histology  Study authors contacted for clarification of study data

**Methodological quality**

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

**Kockanat 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? 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

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

**Kucukyilmaz 2015**
**Study characteristics**

Patient Sampling Method of sampling: selected

**Visual or visual-tactile examination to detect and inform the diagnosis of enamel caries (Review)**

**Kucukyilmaz 2015** (Continued)

	<p>Included conditions: non-cavitated and early lesions</p> <p>Teeth: primary molars pre and post-exfoliation</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: Turkey</p> <p>Setting: in vivo study conducted in dental hospital, followed by in vitro after extraction</p> <p>Number of participants/teeth/sites: 200 teeth</p> <p>Prevalence: enamel = 0.82, dentine = 0.33</p>
Index tests	<p>Category of test: visual - Souza-Zaroni</p> <p>Sequence of test(s): visual, radiograph, DIAGNOdent, ECM completed in vivo and in vitro before sectioning of teeth</p> <p>Examiner training and calibration: yes</p> <p>Teeth cleaning prior to examination: "polishes"</p> <p>Tooth drying prior to examination: yes, "visual inspection of dried teeth"</p> <p>Threshold applied: no caries, incipient lesions in enamel, cavity in enamel, cavity in dentine, possible pulpal involvement</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: index tests then reference standard</p> <p>Training of examiner: not reported</p> <p>Blinding to index test: not reported</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep 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	

**Kucukyilmaz 2015** (Continued)

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			Low concern
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			Low concern
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

**Kucukyilmaz 2015** (Continued)

Were all patients included in the analysis? Yes

**Could the patient flow have introduced bias?** Low risk

**Kuhnisch 2009**
**Study characteristics**

Patient Sampling	Method of sampling: selected Included conditions: sound and non-cavitated Teeth: third molars Sealants: excluded Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Germany Setting: extracted teeth Number of participants/teeth/sites: 54 teeth Prevalence: enamel = 0.89, dentine = 0.35
Index tests	Category of test: visual - Ekstrand - "visually examined with dental magnifying glasses (2x magnification)" Sequence of test(s): visual, radiograph, before sectioning of teeth Examiner training and calibration: 2 examiners independently followed by consensus, training not reported Teeth cleaning prior to examination: cleaning with rotated brushes and air polishing Tooth drying prior to examination: not specifically reported Threshold applied: Ekstrand criteria
Target condition and reference standard(s)	Category: histology - 2 methods light microscope and microradiographic images Sequence of index test and reference standard: index tests then reference standard Training of examiner: 2 examiners, experience unclear Blinding to index test: 1 examiner was involved in the index test evaluation Multiple tests: yes - 2 tests separated by 2 weeks

**Kuhnisch 2009** (Continued)

Site selection: sectioned teeth

Target condition: sound, enamel, outer half of dentine, deep 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

Light microscope used for results

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		

**Kuhnisch 2009** *(Continued)*

**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

**Kuhnisch 2009a**
**Study characteristics**

Patient Sampling	Method of sampling: selected  Included conditions: "sound and mostly non-cavitated", uncertain on inclusion of dentine caries  Teeth: third molars  Sealants: excluded  Restorations: excluded  Surface: occlusal
Patient characteristics and setting	Age: not reported  Sex: not reported  Ethnicity: not reported  Country: Germany  Setting: extracted teeth  Number of participants/teeth/sites: 65 teeth  Prevalence: enamel = 0.82, dentine = 0.35
Index tests	Category of test: visual - UniViSS  Sequence of test(s): visual, prior to reference standard  Examiner training and calibration: 2 trained examiners - independently assessed followed by consensus  Teeth cleaning prior to examination: not reported  Tooth drying prior to examination: not reported

**Kuhnisch 2009a** (Continued)

Threshold applied: UniViSS codes

Target condition and reference standard(s)

Category: histology - light microscope

Sequence of index test and reference standard: visual then histology

Training of examiner: not reported

Blinding to index test: stated, blinded to index test

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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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)?			
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	



**Kuhnisch 2009a** (Continued)

**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

**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

**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)

**Laitala 2017** (Continued)

Index tests

Category of test: visual - ICDAS

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: experienced examiners - trained and calibrated during similar studies

Teeth cleaning prior to examination: not reported

Tooth drying prior to examination: not reported

Threshold applied: ICDAS codes

Target condition and reference standard(s)

Category: bitewing radiograph

Sequence of index test and reference standard: radiograph performed and interpreted separately to 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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		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? 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? 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? Unclear

**Could the patient flow have introduced bias?** Low risk

**Mendes 2006**
**Study characteristics**

Patient Sampling	Method of sampling: selected
	Included conditions: non-cavitated
	Teeth: primary molars
	Sealants: not reported

**Mendes 2006** (Continued)

	Restorations: not reported
	Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 79 teeth, 110 sites Prevalence: enamel = 0.75, dentine = 0.25
Index tests	Category of test: visual - Ekstrand Sequence of test(s): visual, then DIAGNOdent, then radiograph Examiner training and calibration: not reported Teeth cleaning prior to examination: brush, pumice and slurry Tooth drying prior to examination: 5 seconds Threshold applied: Ekstrand criteria, scores 1 & 2 enamel, 3 & 4 dentine
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: not reported Multiple tests: no Site selection: marked on tooth then sectioned Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep 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	
<b>Methodological quality</b>	
<b>Item</b>	<b>Authors' judgement</b>
	<b>Risk of bias</b>
	<b>Applicability concerns</b>

**Mendes 2006** (Continued)

**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	
<b>Could the selection of patients have introduced bias?</b>		High risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Unclear risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Mitropoulos 2010**

**Study characteristics**

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: sound, early/enamel lesions, no cavitation (no distinct enamel breakdown on either one or both of their proximal surfaces)</p> <p>Teeth: molars and premolars</p> <p>Sealants: not reported</p> <p>Restorations: excluded</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Greece</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 20 teeth, 40 sites</p> <p>Prevalence: enamel = 0.60, dentine = 0.45</p>
Index tests	<p>Category of test: visual - ICDAS, not placed in a model to recreate the approximal contact</p> <p>Sequence of test(s): 1) visual, 2) radiography and histology (precise sequence not stated - but classified blindly and with a 10-day break between viewings for each modality)</p> <p>Examiner training and calibration: 2 experienced dentists with an interest in cariology. Calibration performed through a prior pilot study</p> <p>Teeth cleaning prior to examination: "cleaned with toothbrush and water"</p> <p>Tooth drying prior to examination: 3-in-1 air syringe</p> <p>Threshold applied: ICDAS - codes</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: as above (see index test)</p> <p>Training of examiner: as above (see index test)</p> <p>Blinding to index test: yes - although not clear if reference standard was last, they were "classified blindly... with a 10-day intermission among the different modalities"</p> <p>Multiple tests: no</p> <p>Site selection: "longitudinally bisected... in the centre of the suspected carious lesion"</p> <p>Target condition: 0 = no lesion, E1 = caries extended to outer half of the enamel, E2 = caries extending to the inner half of the enamel, D1 = caries extending to dentine</p>

**Mitropoulos 2010** (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 Data calculated from crosstabs; used the data for examiner 1

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	

**Mitropoulos 2010** *(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

**Nakagawa 2013**
**Study characteristics**

Patient Sampling	Method of sampling: "arbitrarily chosen for investigation"  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  Teeth: not reported  Sealants: not reported Restorations: not reported  Surface: smooth
Patient characteristics and setting	Age: not reported  Sex: not reported  Ethnicity: not reported  Country: Japan  Setting: extracted teeth  Number of participants/teeth/sites: 93 teeth, 127 sites  Prevalence: enamel = 0.75, dentine = 0.31
Index tests	Category of test: visual - scale no specified  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"  Examiner training and calibration: 2 clinical experience of over 9 years and 2 with 3 years



**Nakagawa 2013** (Continued)

Teeth cleaning prior to examination: "prophylaxis paste using a brush cone attached to a low-speed handpiece"

Tooth drying prior to examination: dried

Threshold applied: sound, enamel demineralization without surface breakdown, enamel breakdown due to caries, dentine caries

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 demineralization, 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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

**Nakagawa 2013** (Continued)

If a threshold was used, was it pre-specified?	Yes	
If multiple tests were applied were different examiners used for each (in vivo)?	No	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		High
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Nakajima 2014**
**Study characteristics**

Patient Sampling	Method of sampling: "Twenty-six primary teeth were selected"  Included conditions: sound and early lesions, "38 investigation sites of occlusal fissures (noncavitated and cavitated) were selected" - level of cavitation uncertain  Teeth: primary molars, "Extracted human primary molar teeth with/without occlusal caries"  Sealants: not reported Restorations: not reported  Surface: occlusal
------------------	--

**Nakajima 2014** (Continued)

Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported 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: visual - scale not specified - viewed on a monitor following digital imaging 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, enamel demineralization but without cavitation, localised enamel breakdown, dentine caries		
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 demineralization, 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	Mean results used from 6 examiners		
<b>Methodological quality</b>			
<b>Item</b>	<b>Authors' judgement</b>	<b>Risk of bias</b>	<b>Applicability concerns</b>

**DOMAIN 1: Patient Selection**

**Nakajima 2014** (Continued)

Was a consecutive or random sample of patients enrolled?	No	
Was a case-control design avoided?	Yes	
Did the study avoid inappropriate exclusions?	Unclear	
<b>Could the selection of patients have introduced bias?</b>		High risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		High
<b>DOMAIN 2: Index Test (All tests)</b>		
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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		High
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

## Neuhaus 2011

### Study characteristics

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: unclear</p> <p>Teeth: primary molars (first and second)</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: Switzerland</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 37 teeth, 37 sites</p> <p>Prevalence: enamel = 0.73, dentine = 0.24</p>
Index tests	<p>Category of test: visual - ICDAS</p> <p>Sequence of test(s): index tests (visual, DIAGNOdent devices then radiograph) then reference standard</p> <p>Examiner training and calibration: experienced examiners</p> <p>Teeth cleaning prior to examination: 3-in-1 syringe</p> <p>Tooth drying prior to examination: not reported</p> <p>Threshold applied: ICDAS codes</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: not reported</p> <p>Multiple tests: no</p> <p>Site selection: unclear</p> <p>Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep 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>

**Neuhaus 2011** (Continued)

Participants receiving both tests but excluded from results: 0

Comparative

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		High risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			

**Neuhaus 2011** (Continued)

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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Neuhaus 2015**
**Study characteristics**

Patient Sampling	Method of sampling: selected Included conditions: unclear Teeth: premolars and molars Sealants: unclear Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 100 teeth Prevalence: enamel = 0.87, dentine = 0.24
Index tests	Category of test: visual - ICDAS, plus magnification with loupe Sequence of test(s): index tests (visual, with 3 different loupe systems) then reference standard Examiner training and calibration: 3 groups (3rd year students, 4th year students, dentists 2 to 15 years post-graduation), all calibrated Teeth cleaning prior to examination: cleaned using "air abrasion with sodium bicarbonate for 15 s and rinsed with water-air-spray for 10 s" Tooth drying prior to examination: not reported Threshold applied: ICDAS codes
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: 2 experienced examiners

**Neuhaus 2015** (Continued)

Blinding to index test: not reported

Multiple tests: no

Site selection: sectioned teeth

Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep 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 Data taken for dentists and computed for 98 teeth

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
Is the reference standards likely to correctly classify the target condition?	Yes		



**Neuhaus 2015** (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Novaes 2009**

<b>Study characteristics</b>	
Patient Sampling	Method of sampling: random Included conditions: no cavitation and early lesions Teeth: primary molars (first and second) Sealants: unclear Restorations: excluded Surface: approximal
Patient characteristics and setting	Age: 5 to 12 years, mean 7.7 Sex: 21 male, 29 female Ethnicity: not reported Country: Brazil Setting: clinical setting dental hospital Number of participants/teeth/sites: 50 participants, 621 sites Prevalence: enamel = 0.41, dentine = 0.03
Index tests	Category of test: visual - ICDAS, no model created to replicate approximal contact Sequence of test(s): index tests (visual, radiograph, DIAGNOdent pen) then reference standard Examiner training and calibration: trained but no calibration Teeth cleaning prior to examination: brush and slurry

**Novaes 2009** (Continued)

	Tooth drying prior to examination: air dried, 5 seconds Threshold applied: ICDAS codes
Target condition and reference standard(s)	Category: visual - separators 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: approximal surface Target condition: sound, white spot, cavitated
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: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0

Comparative

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		Low risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			Low concern
<b>DOMAIN 2: Index Test (All tests)</b>			
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		

**Novaes 2009** (Continued)

<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	Low concern
<b>DOMAIN 3: Reference Standard</b>	
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
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	High risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Novaes 2010**

<b>Study characteristics</b>	
Patient Sampling	Method of sampling: randomly selected from children seeking dental treatment  Included conditions: no cavitation and early lesions, frank and severe lesions excluded  Teeth: primary molars (first and second), in contact with another tooth  Sealants: unclear  Restorations: excluded  Surface: approximal
Patient characteristics and setting	Age: 4 to 12 years, mean 7.25  Sex: 32 male, 44 female  Ethnicity: not reported  Country: Brazil  Setting: clinical setting dental hospital

**Novaes 2010** (Continued)

	Number of participants/teeth/sites: 76 participants, 592 sites Prevalence: enamel = 0.81, dentine = 0.05
Index tests	Category of test: visual - ICDAS Sequence of test(s): index tests (visual, radiograph, DIAGNOdent pen) then reference standard Examiner training and calibration: experienced and trained but no calibration Teeth cleaning prior to examination: "cleaned with a rotating bristle brush with pumice/water slurry and dental floss" Tooth drying prior to examination: air dried, 5 seconds Threshold applied: ICDAS codes
Target condition and reference standard(s)	Category: visual - separators 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: approximal surface Target condition: sound, white spot, cavitated
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: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

**Methodological quality**

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

**Novaes 2010** (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? 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? 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

**Novaes 2012**
**Study characteristics**

Patient Sampling	Method of sampling: selected Included conditions: unclear Teeth: primary molars - "recently extracted primary molars were selected" Sealants: excluded
------------------	---

**Novaes 2012** (Continued)

	Restorations: excluded
	Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Sao Paulo, Brazil Setting: extracted teeth Number of participants/teeth/sites: 77 teeth, 113 sites: "One or two suspect sites on each occlusal surface were selected" Prevalence: enamel = 0.57, dentine = 0.17
Index tests	Category of test: visual - ICDAS Sequence of test(s): index tests (radiograph,visual, DIAGNOdent, VistaProof) then reference standard Examiner training and calibration: trained but no calibration Teeth cleaning prior to examination: yes Tooth drying prior to examination: yes Threshold applied: ICDAS codes
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 Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep 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: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	
<b>Methodological quality</b>	

**Novaes 2012** (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			
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		
<b>Could the patient flow have introduced bias?</b>		Low risk	

## Paula 2011

### Study characteristics

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: no cavitation and early lesions</p> <p>Teeth: permanent third 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 - tooth bank</p> <p>Number of participants/teeth/sites: 26 teeth, 64 sites</p> <p>Prevalence: enamel = 0.88, dentine = 0.28</p>
Index tests	<p>Category of test: visual - no specific scale</p> <p>Sequence of test(s): index tests followed by reference standard</p> <p>Examiner training and calibration: experienced</p> <p>Teeth cleaning prior to examination: pumice slurry and water</p> <p>Tooth drying prior to examination: air dried</p> <p>Threshold applied: 0 = no caries, 1 = carious lesion in the enamel, 2 = carious lesion in the dentine</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: 0</p> <p>Participants with reference standard but no index test: 0</p> <p>Time interval between tests: 1 to 2 days</p>



Paula 2011 (Continued)

Participants receiving both tests but excluded from results: 0

Comparative

Notes

Data taken from in vitro element of study

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			

**Paula 2011** (Continued)

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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Pereira 2011**
**Study characteristics**

Patient Sampling	Method of sampling: selected, "None of the teeth showed macroscopic signs of cavity formation" Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: unclear Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 96 teeth Prevalence: enamel = 0.57, dentine = 0.25
Index tests	Category of test: visual - Ekstrand Sequence of test(s): index tests (visual, radiograph, ECM, DIAGNOdent, QLF) then reference standard Examiner training and calibration: training event Teeth cleaning prior to examination: paste and rotating brush Tooth drying prior to examination: yes Threshold applied: Ekstrand criteria
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: "Three examiners underwent a training session, which consisted of 2 h of theoretical training and 4 h of practice on extracted teeth"

**Pereira 2011** (Continued)

Blinding to index test: not reported

Multiple tests: no

Site selection: sectioned teeth

Target condition: no caries; demineralization extending to the outer half of the enamel; demineralization extending to the inner half of the enamel; demineralization extending to the outer half of the dentine; demineralization extending to the outer half of the 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: 1 week to allow for separation of teeth

Participants receiving both tests but excluded from results: 0

Comparative

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			

**Pereira 2011** (Continued)

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
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern
<b>DOMAIN 4: Flow and Timing</b>	
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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Qudeimat 2019**
**Study characteristics**

Patient Sampling	Method of sampling: selected from a pool of 651 teeth Included conditions: unclear Teeth: permanent molars and premolars Sealants: excluded Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Kuwait Setting: extracted teeth Number of participants/teeth/sites: 270 teeth Prevalence: enamel = 0.90, dentine = 0.28
Index tests	Category of test: visual - ICDAS Sequence of test(s): index tests then reference standard

**Qudeimat 2019** (Continued)

	Examiner training and calibration: calibration and training completed over a 1-week period, multiple experienced examiners  Teeth cleaning prior to examination: yes - no details  Tooth drying prior to examination: 3-in-1 syringe available  Threshold applied: ICDAS criteria
Target condition and reference standard(s)	Category: histology  Sequence of index test and reference standard: index test then reference standard  Training of examiner: "prior experience in caries histological classification"  Blinding to index test: not reported  Multiple tests: no  Site selection: sectioned teeth  Target condition: no caries, outer/inner enamel, outer/inner 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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			Low concern
<b>DOMAIN 2: Index Test (All tests)</b>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

**Qudeimat 2019** (Continued)

If a threshold was used, was it pre-specified?	Yes	
If multiple tests were applied were different examiners used for each (in vivo)?		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		High
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Rocha 2003**
**Study characteristics**

Patient Sampling	Method of sampling: selected, "Fifty occlusal sites were selected for this study" Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: excluded Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: 10 to 11 years Sex: not reported

**Rocha 2003** (Continued)

	Ethnicity: not reported
	Country: Brazil
	Setting: clinic based, children with teeth close to exfoliation
	Number of participants/teeth/sites: 29 participants, 50 sites
	Prevalence: enamel = 0.58, dentine = 0.14
Index tests	Category of test: visual - Ekstrand  Sequence of test(s): DIAGNOdent, visual, radiograph followed by reference standard - visual could have been influenced by other tests  Examiner training and calibration: previously trained, "independently by 2 previously trained examiners"  Teeth cleaning prior to examination: not in clinical setting  Tooth drying prior to examination: air/water spray  Threshold applied: Ekstrand criteria
Target condition and reference standard(s)	Category: histology  Sequence of index test and reference standard: index test conducted before reference standard  Training of examiner: not reported  Blinding to index test: yes  Multiple tests: no  Site selection: sectioned tooth  Target condition: sound, outer enamel, inner enamel and first third dentine middle and inner 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	
<b>Methodological quality</b>	
<b>Item</b>	<b>Authors' judgement</b> <b>Risk of bias</b> <b>Applicability concerns</b>
<b>DOMAIN 1: Patient Selection</b>	
Was a consecutive or random sample of patients enrolled?	No
Was a case-control design avoided?	Yes

**Rocha 2003** (Continued)

Did the study avoid inappropriate exclusions?	Yes	
<b>Could the selection of patients have introduced bias?</b>		High risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		Low concern
<b>DOMAIN 2: Index Test (All tests)</b>		
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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		High risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		Low concern
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Rodrigues 2008**
**Study characteristics**

Patient Sampling	Method of sampling: selected
	Included conditions: no cavitation and early lesions



**Rodrigues 2008** (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: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 119 teeth Prevalence: enamel = 0.93, dentine = 0.54
Index tests	Category of test: visual - ICDAS Sequence of test(s): DIAGNOdent, DIAGNOdent pen, VistaProof, visual, radiograph Examiner training and calibration: experienced Teeth cleaning prior to examination: yes Tooth drying prior to examination: yes Threshold applied: ICDAS codes
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests conducted before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, outer dentine, inner 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	
<b>Methodological quality</b>	

**Rodrigues 2008** (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			
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		
<b>Could the patient flow have introduced bias?</b>		Low risk	

## Rodrigues 2009

### Study characteristics

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: no cavitation and early lesions</p> <p>Teeth: primary 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: 148 teeth</p> <p>Prevalence: enamel = 0.92, dentine = 0.03</p>
Index tests	<p>Category of test: visual - not specified</p> <p>Sequence of test(s): visual and DIAGNOdent combined in 1 examination</p> <p>Examiner training and calibration: calibrated</p> <p>Teeth cleaning prior to examination: yes</p> <p>Tooth drying prior to examination: yes</p> <p>Threshold applied: sound, enamel, outer/inner dentine</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: index tests conducted before 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, outer dentine, inner 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>

**Rodrigues 2009** (Continued)

Participants receiving both tests but excluded from results: 0

Comparative

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			

**Rodrigues 2009** (Continued)

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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Rodrigues 2009a**
**Study characteristics**

Patient Sampling	Same study as <a href="#">Rodrigues 2009</a> but entered for analysis of permanent teeth
Patient characteristics and setting	
Index tests	
Target condition and reference standard(s)	
Flow and timing	
Comparative	
Notes	

**Rodrigues 2013**
**Study characteristics**

Patient Sampling	Method of sampling: not reported Included conditions: not reported Teeth: permanent first second and third molars Sealants: excluded Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 170 teeth

**Rodrigues 2013** (Continued)

	Prevalence: enamel = 0.83, dentine = 0.46
Index tests	Category of test: visual - ICDAS (before and after e-learning program) Sequence of test(s): visual prior to reference standard Examiner training and calibration: "four dentists in Switzerland with an average of eight years (range seven to nine years)" Teeth cleaning prior to examination: calculus and debris were removed using a scaler, then rinsed with water and brush for 15 seconds, then water jet cleaned with sodium hydrogen carbonate Tooth drying prior to examination: yes, 3-in-1 syringe Threshold applied: ICDAS criteria
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests conducted before reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer/inner enamel, outer/inner 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	Data taken from results after the training programme

**Methodological quality**

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

**Rodrigues 2013** (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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Seremidi 2012**
**Study characteristics**

Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars Sealants: excluded Restorations: excluded Surface: occlusal
------------------	---

**Seremidi 2012** (Continued)

Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Greece Setting: extracted teeth Number of participants/teeth/sites: 41 teeth, 107 sites Prevalence: enamel = 0.78, dentine = 0.19		
Index tests	Category of test: visual - Ekstrand Sequence of test(s): index tests (visual followed by DIAGNOdent pen and VistaProof) conducted prior to reference standard Examiner training and calibration: 1 calibrated examiner Teeth cleaning prior to examination: yes - rubber cup and air water syringe Tooth drying prior to examination: 5 seconds compressed air Threshold applied: no caries, opacity or discolouration visible after air drying, opacity or discolouration visible without air drying, grayish discolouration from the underlying dentine, cavitation exposing the dentine beneath		
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests conducted before reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer enamel, inner enamel, outer dentine, inner 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			
<b>Methodological quality</b>			
<b>Item</b>	<b>Authors' judgement</b>	<b>Risk of bias</b>	<b>Applicability concerns</b>



**Seremidi 2012** (Continued)

**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	
<b>Could the selection of patients have introduced bias?</b>		High risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Shi 2000**

**Study characteristics**

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: not clearly reported, "selected for the study if a first permanent molar was erupted"</p> <p>Teeth: permanent molars and premolars</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: Sweden</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 76 teeth/surfaces</p> <p>Prevalence: enamel = 0.73, dentine = 0.39</p>
Index tests	<p>Category of test: visual - no scale described</p> <p>Sequence of test(s): index then reference standard</p> <p>Examiner training and calibration: not reported</p> <p>Teeth cleaning prior to examination: yes, technique not reported</p> <p>Tooth drying prior to examination: yes</p> <p>Threshold applied: sound, dentine caries or caries with no-cavitation</p> <p>Note: dentine caries were not reported in the analysis</p>
Target condition and reference standard(s)	<p>Category: histology with microradiograph</p> <p>Sequence of index test and reference standard: index test before reference standard</p> <p>Training of examiner: not reported</p> <p>Blinding to index test: no</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth according to photographed locations</p> <p>Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine</p>
Flow and timing	<p>Participants with index test but no reference standard: 6</p> <p>Participants with reference standard but no index test: 0</p>

**Shi 2000** (Continued)

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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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)?			
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		High risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			

**Shi 2000** (Continued)

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
<b>Could the patient flow have introduced bias?</b>	<b>High risk</b>

**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 are sampled  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: visual - Ekstrand  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  Teeth cleaning prior to examination: "cleaned using a brush cone attached to low speed handpiece with prophylaxis paste"  Tooth drying prior to examination: yes  Threshold applied: Ekstrand criteria
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

**Shimada 2010** (Continued)

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 demineralization, 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
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High

**Shimada 2010** (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

**Sidi 1988**
**Study characteristics**

Patient Sampling	Method of sampling: selected - participants taking part in a caries trial Included conditions: only surfaces considered to be sound by clinical evaluation were evaluated Teeth: permanent molar and premolars Sealants: not reported Restorations: not reported Surface: approximal
Patient characteristics and setting	Age: 12 to 13 years Sex: not reported Ethnicity: not reported Country: not reported Setting: clinical trial "carried out away from the dental surgery" Number of participants/teeth/sites: 456 participants, 4405 surfaces (71 participants could not be radiographed within the study) Prevalence: enamel = 0.03, dentine = 0.01 (very low prevalence results in high concern for applicability)

**Sidi 1988** (Continued)

Index tests	Category of test: visual - scale not reported Sequence of test(s): visual examination preceded radiographs and transillumination, radiographs developed and interpreted at a later date Examiner training and calibration: not reported Teeth cleaning prior to examination: not reported Tooth drying prior to examination: compressed air available for drying Threshold applied: sound, enamel, dentine
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 - attempt made to blind examiners to index test results Multiple tests: visual completed but without separation so not a viable reference standard Site selection: all surfaces Target condition: sound, enamel, or dentine
Flow and timing	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 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	Radiograph treated as reference standard for these results and transillumination, so not included in imaging modalities review

**Methodological quality**

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

**Sidi 1988** (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?

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?**

Unclear

**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

**Souza 2013**
**Study characteristics**

Patient Sampling

Method of sampling: unclear



**Souza 2013** (Continued)

	<p>Included conditions: "occlusal surfaces varying from sound to having different stages of carious lesions"</p> <p>Teeth: primary 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: Brazil</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 79 teeth (42 first molars and 37 second molars)</p> <p>Prevalence: enamel = 0.76, dentine = 0.35</p>
Index tests	<p>Category of test: visual - ICDAS</p> <p>Sequence of test(s): visual, radiograph, DIAGNOdent, DIAGNOdent pen and VistaProof, then reference standard; "teeth were mounted individually on a dental model"</p> <p>Examiner training and calibration: "Two experienced examiners independently assessed the teeth"</p> <p>Teeth cleaning prior to examination: yes, with sodium bicarbonate and water-powder blasting device for 10 seconds</p> <p>Tooth drying prior to examination: first dry and then after 5 seconds air drying</p> <p>Threshold applied: ICDAS codes</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: index test before reference standard</p> <p>Training of examiner: "experienced senior researcher, who did not take part in the examination"</p> <p>Blinding to index test: not reported</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: 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>

Souza 2013 (Continued)

Participants receiving both tests but excluded from results: 0

Comparative

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			

**Souza 2013** (Continued)

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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Soviero 2012**
**Study characteristics**

Patient Sampling	Method of sampling: selected  Included conditions: "no enamel developmental defects", "proximal surfaces with extensive cavities involving more than half of the proximal surface were excluded"  Teeth: primary molars  Sealants: not reported  Restorations: excluded  Surface: approximal
Patient characteristics and setting	Age: not reported  Sex: not reported  Ethnicity: not reported  Country: Brazil  Setting: extracted teeth  Number of participants/teeth/sites: 25 teeth, 48 surfaces  Prevalence:enamel = 0.81, dentine = 0.33
Index tests	Category of test: visual - ICDAS, no model created to replicate approximal contact  Sequence of test(s): index tests (visual, radiograph, microCT) then reference standard  Examiner training and calibration: experienced in clinical and radiographic caries diagnosis  Teeth cleaning prior to examination: extracted teeth - not reported  Tooth drying prior to examination: "Teeth were examined wet and after 5 s of air-drying"  Threshold applied: ICDAS codes
Target condition and reference standard(s)	Category: histology

**Soviero 2012** (Continued)

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

Training of examiner: experienced examiners

Blinding to index test: unclear

Multiple tests: no

Site selection: approximal surface

Target condition: sound, outer/inner enamel, outer/inner 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: "from two to three weeks"

Participants receiving both tests but excluded from results: 0

Comparative

Notes

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High

**Soviero 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?	Unclear
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	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
<b>Could the patient flow have introduced bias?</b>	Low risk

**Sridhar 2009**
**Study characteristics**

Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent molars and premolars Sealants: excluded Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: India Setting: extracted teeth Number of participants/teeth/sites: 50 teeth Prevalence: enamel = 0.96, dentine = 0.12
Index tests	Category of test: visual - Ekstrand Sequence of test(s): index tests (visual, radiograph, DIAGNOdent) then reference standard

**Sridhar 2009** (Continued)

	Examiner training and calibration: not reported Teeth cleaning prior to examination: scaled with paste Tooth drying prior to examination: air dried Threshold applied: Ekstrand criteria
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test before reference standard Training of examiner: not reported Blinding to index test: no Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer enamel, inner enamel, outer dentine, inner 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: 2
Comparative	
Notes	

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			Low concern
<b>DOMAIN 2: Index Test (All tests)</b>			
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		

**Sridhar 2009** (Continued)

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?** 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

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

**Teo 2014**
**Study characteristics**

Patient Sampling	Method of sampling: unclear Included conditions: no cavitation and early lesions Teeth: primary molars - selected for extraction Sealants: excluded Restorations: excluded Surface: occlusal
Patient characteristics and setting	Age: 2 to 11 years Sex: not reported Ethnicity: not reported Country: UK

**Teo 2014** (Continued)

	Setting: dental school (in vivo study used, but in vitro also available)
	Number of participants/teeth/sites: 64 teeth/surfaces
	Prevalence: enamel = 0.72, dentine = 0.31
Index tests	Category of test: visual - ICDAS  Sequence of test(s): index tests (visual, DIAGNOdent pen, CarieScan PRO) then reference standard  Examiner training and calibration: yes on subsample  Teeth cleaning prior to examination: pumice and a bristle brush  Tooth drying prior to examination: assessed before and after drying  Threshold applied: ICDAS codes
Target condition and reference standard(s)	Category: histology  Sequence of index test and reference standard: index test before reference standard  Training of examiner: not reported  Blinding to index test: no  Multiple tests: no  Site selection: recorded on a drawing of the occlusal surface  Target condition: sound, outer enamel, inner enamel, outer dentine, inner 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	
<b>Methodological quality</b>	
<b>Item</b>	<b>Authors' judgement</b>
	<b>Risk of bias</b>
	<b>Applicability concerns</b>
<b>DOMAIN 1: Patient Selection</b>	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes



**Teo 2014** (Continued)

<b>Could the selection of patients have introduced bias?</b>		Unclear risk
<b>Are there concerns that the included patients and setting do not match the review question?</b>		Low concern
<b>DOMAIN 2: Index Test (All tests)</b>		
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	
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>		Low concern
<b>DOMAIN 3: Reference Standard</b>		
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	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>		Low concern
<b>DOMAIN 4: Flow and Timing</b>		
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	
<b>Could the patient flow have introduced bias?</b>		Low risk

**Tonkaboni 2019**

<b>Study characteristics</b>		
Patient Sampling		Method of sampling: not reported
		Included conditions: "Teeth with large proximal cavitated carious lesions with extensive tooth destruction were excluded and replaced"
		Teeth: permanent molars and premolars

**Tonkaboni 2019** (Continued)

	<p>Sealants: not reported</p> <p>Restorations: not reported</p> <p>Surface: approximal</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>Ethnicity: not reported</p> <p>Country: Iran</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 108 teeth, evaluated in 3 regions (a total of 324 regions)</p> <p>Prevalence: contact area and higher - enamel = 0.42, dentine = 0.35</p>
Index tests	<p>Category of test: visual - ICDAS, "mounted in putty impression material next to each other such that they were in contact at their marginal ridges"</p> <p>Sequence of test(s): index tests (visual, radiograph, VistaCam) then reference standard</p> <p>Examiner training and calibration: not reported</p> <p>Teeth cleaning prior to examination: brushed and scaled</p> <p>Tooth drying prior to examination: yes but technique not reported</p> <p>Threshold applied: ICDAS codes</p>
Target condition and reference standard(s)	<p>Category: histology</p> <p>Sequence of index test and reference standard: index test before reference standard</p> <p>Training of examiner: not reported</p> <p>Blinding to index test: no</p> <p>Multiple tests: no</p> <p>Site selection: sectioned teeth</p> <p>Target condition: sound, outer enamel, inner enamel, outer dentine, inner 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	Data used from results of site at the contact area or higher

**Methodological quality**

**Tonkaboni 2019** (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Unclear risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern
<b>DOMAIN 4: Flow and Timing</b>			
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		

**Tonkaboni 2019** (Continued)

**Could the patient flow have introduced bias?**

Low risk

**Xiao-Hua 2016**

**Study characteristics**

Patient Sampling	<p>Method of sampling: selected</p> <p>Included conditions: sound and non-cavitated</p> <p>Teeth: "permanent posterior teeth"</p> <p>Sealants: not reported</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: China</p> <p>Setting: extracted teeth</p> <p>Number of participants/teeth/sites: 77 teeth (97 reported in results)</p> <p>Prevalence: enamel = 0.74, dentine = 0.48</p>
Index tests	<p>Category of test: visual - Ekstrand</p> <p>Sequence of test(s): visual and OCT, all prior to reference standard</p> <p>Examiner training and calibration: 3 examiners reached a consensus, training unclear</p> <p>Teeth cleaning prior to examination: brushed and scaled</p> <p>Tooth drying prior to examination: yes but technique not reported</p> <p>Threshold: Ekstrand criteria</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: 0</p>

**Xiao-Hua 2016** (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

Need formal translation details taken from Google translate

**Methodological quality**

Item	Authors' judgement	Risk of bias	Applicability concerns
<b>DOMAIN 1: Patient Selection</b>			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<b>Could the selection of patients have introduced bias?</b>		High risk	
<b>Are there concerns that the included patients and setting do not match the review question?</b>			High
<b>DOMAIN 2: Index Test (All tests)</b>			
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		
<b>Could the conduct or interpretation of the index test have introduced bias?</b>		Low risk	
<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>			High
<b>DOMAIN 3: Reference Standard</b>			
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		
<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>		Low risk	
<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>			Low concern

**Xiao-Hua 2016** (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?	Yes
<b>Could the patient flow have introduced bias?</b>	Low risk

BW = bitewing; DIFOTI = digital imaging fibre optic transillumination; DTA = diagnostic test accuracy; ECM = electronic caries monitor; FOTI = fibre optic transillumination; ICDAS = International Caries Detection and Assessment System; microCT = microcomputed tomography; n/a = not applicable; 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
<a href="#">Abrams 2017</a>	No clear results of reference standard, cannot construct a 2 x 2 table
<a href="#">Almosa 2014</a>	Visual used as reference standard for an index test of fluorescence
<a href="#">Amaechi 2013</a>	Uses index test to inform "ground truth" so no valid reference standard
<a href="#">Apostolopoulou 2009</a>	Not possible to calculate 2 x 2 table for visual test from data presented, table 1 confirms 2 sites are sound from histology, but specificity in table 2 for direct visual is 0.21
<a href="#">Ashley 2000</a>	Cannot correctly construct a 2 x 2 table
<a href="#">Askaroglou 2011</a>	Not a DTA study, investigates correlation effects of sealants on fluorescence results
<a href="#">Attrill 2001</a>	"Data for the detection of enamel caries were not calculated because of the small numbers of teeth with enamel caries"
<a href="#">Bengtson 2005</a>	Sound and initial enamel caries lesions were combined as the disease negative result
<a href="#">Bizhang 2016</a>	Data not available for 2 x 2 table
<a href="#">Bozdemir 2013</a>	Data not useable, only excavated teeth reported
<a href="#">Braga 2009a</a>	".. occlusal surfaces represented a wide range of caries appearances. Teeth with large cavities were excluded" therefore teeth with cavities were intentionally included in the sample
<a href="#">Chen 2012</a>	Inappropriate reference standard, unable to validate the visual index test
<a href="#">Chong 2003</a>	Considered as visual reference standard in fluorescence and imaging modalities reviews so cannot include here
<a href="#">Chu 2010</a>	Included cavitated teeth
<a href="#">Cortes 2003</a>	Included cavitated teeth
<a href="#">Costa 2007</a>	Not possible to extract full 2 x 2 table

Study	Reason for exclusion
<a href="#">Diniz 2011a</a>	Visual inspected but not assessed
<a href="#">Dong 2007</a>	Cannot construct a 2 x 2 table
<a href="#">Duruturk 2011</a>	Visual and radiograph combined to act as reference standard
<a href="#">El-Damanhoury 2014</a>	"... extracted teeth representing ICDAS II scores between 0 and 6 on their occlusal surfaces"
<a href="#">Elhennawy 2018</a>	Selected ICDAS lesions from 0 to 6, therefore dentinal caries included
<a href="#">Erten 2006</a>	Sensitivity and specificity not presented, 2 x 2 table not possible
<a href="#">Forgie 2002</a>	Cannot complete 2 x 2 table. 100 teeth sampled, 138 caries lesions detected, but no report of sound surfaces/teeth to generate a prevalence
<a href="#">Fyffe 2000</a>	Dentinal caries included
<a href="#">Goel 2016</a>	Cannot construct a 2 x 2 table
<a href="#">Gomez 2013</a>	Dentinal caries included
<a href="#">Graye 2012</a>	Cannot construct a 2 x 2 table
<a href="#">Heinrich-Weltzien 2003</a>	Visual examination used as part of the composite reference standard, those that were identified visually as dentine care were excavated, therefore cannot treat visual as an index test
<a href="#">Huysmans 1998</a>	Dentinal caries included
<a href="#">Iranzo-Cortés 2018</a>	Same study as <a href="#">Iranzo-Cortes 2017</a>
<a href="#">Jablonski-Momeni 2009</a>	Cannot completed 2 x 2 table as prevalence of disease not reported
<a href="#">Jablonski-Momeni 2013</a>	Recruitment strategy aimed to recruit dentinal lesions
<a href="#">Jablonski-Momeni 2018</a>	Included ICDAS 6 lesions - extreme cavitation
<a href="#">Kavvadia 2008</a>	Cannot complete 2 x 2 table due to unclear reporting
<a href="#">Kavvadia 2012</a>	Cannot complete 2 x 2 table due to unclear reporting
<a href="#">Khalaf 2018</a>	"Teeth with frank cavitation (ICDAS scores 5 and 6) were also excluded from the study", therefore teeth with ICDAS 4 were included which equate to dentine caries
<a href="#">Kidd 2003</a>	Dentinal caries included
<a href="#">Kordic 2003</a>	Dentinal caries included
<a href="#">Kuhnisch 2011</a>	Same study as <a href="#">Kuhnisch 2009a</a> , not possible to create 2 x 2 table from the results of the 2011 paper
<a href="#">Lussi 1991</a>	Cut-off point for sensitivity and specificity places "subsurface lesions" with sound which does not allow us to create 2 x 2 table at the correct threshold
<a href="#">Lussi 2003</a>	Data presented at D <sub>2</sub> threshold only

Study	Reason for exclusion
Melo 2017	Teeth (molars and premolars) that had been previously diagnosed for filling, therefore dentinal caries
Mialhe 2011	Sound teeth excluded from the sample, "Those surfaces that were considered sound by all the methods were excluded from the study sample"
Mitropoulos 2012	Cannot construct 2 x 2 table, prevalence of sound, enamel, dentine not reported
Neuhaas 2015a	Dentinal caries included
Novaes 2012a	Cannot construct 2 x 2 table
Pereira 2009	Cannot construct 2 x 2 table
Piovesan 2013	Dentinal caries included
Qudeimat 2016	"Teeth represented a wide range of clinical appearances from apparently sound to cavitated dentine caries"
Reis 2006	Cannot construct 2 x 2 table at D <sub>1</sub> threshold, results reported at D <sub>2</sub>
Ricketts 1995	Dentinal caries included
Rodrigues 2008a	Dentinal caries included
Senel 2010	Cannot construct 2 x 2 table
Shoaib 2009	Dentinal caries included
Silva 2008	Included cavitated teeth
Subka 2019	The sample included teeth with advanced caries: "The teeth extracted usually had advanced caries in one proximal surface but no or less advanced caries on the other"
Umemori 2010	Not a relevant index test: "digital photographs obtained were processed and analysed using image analysis software"
Ünal 2019	Authors confirmed there were no sound teeth in sample so cannot include
Virajsilp 2005	Not possible to construct a 2 x 2 table for visual
Wenzel 1990	Dentinal caries included
White 1978	Unable to construct a 2 x 2 table
Zandona 2009	Dentinal caries included - all ICDAS scores

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

### Characteristics of studies awaiting classification *[ordered by study ID]*



**Luczaj-Cepowicz 2019**

Patient Sampling	In vitro study on extracted teeth. Teeth were selected for the accuracy study or calibration exercise, method of selection not stated
Patient characteristics and setting	Occlusal surfaces of 48 extracted teeth with "no signs of caries, noticeable opacity, or white/brown discoloration that did not match the appearance of healthy enamel." Exclusion criteria were "the presence of fissure sealants, fillings, hypomineralization, occlusoproximal lesions, or lesions with cavitation"
Index tests	ICDAS II
Target condition and reference standard(s)	Enamel threshold all lesions with radiological E1, E2, D1, and D2 progression were considered positive, while 0 was assumed to be sound  Dentine lesions (dentine threshold), radiological progression of D1 and D2 was considered positive, while 0, E1, and E2 were assumed to be sound  Reference standard was microCT
Flow and timing	Time interval between tests not explicitly stated
Comparative	-
Notes	2 dentists undertook examinations. It is unclear whether the sensitivity values refer to a single examiner or are consensus results. Also unclear whether assessments were made on the extracted teeth or on digital images of the teeth

ICDAS = International Caries Detection and Assessment System; microCT = microcomputed tomography.

**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	70	19590
2 ICDAS	38	6817
3 ERK	15	1306
4 Other visual	18	11467

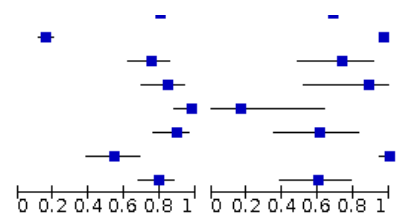
**Test 1. All**

All

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	29	2	7	0	0.81 [0.64, 0.92]	0.00 [0.00, 0.84]		
Akarsu 2006	115	4	12	34	0.91 [0.84, 0.95]	0.89 [0.75, 0.97]		
Ashley 1998	37	11	25	30	0.60 [0.46, 0.72]	0.73 [0.57, 0.86]		
Bahrololoomi 2015	95	3	7	4	0.93 [0.86, 0.97]	0.57 [0.18, 0.90]		
Bottenberg 2016	65	15	4	12	0.94 [0.86, 0.98]	0.44 [0.25, 0.65]		
Braga 2009	73	8	9	41	0.89 [0.80, 0.95]	0.84 [0.70, 0.93]		
Braga 2010	22	5	4	19	0.85 [0.65, 0.96]	0.79 [0.58, 0.93]		
Braga 2010	22	4	4	20	0.85 [0.65, 0.96]	0.83 [0.63, 0.95]		
Braun 2017	64	9	1	10	0.98 [0.92, 1.00]	0.53 [0.29, 0.76]		
Bussanelli 2015	40	3	2	14	0.95 [0.84, 0.99]	0.82 [0.57, 0.96]		
Bussanelli 2015a	43	15	6	13	0.88 [0.75, 0.95]	0.46 [0.28, 0.66]		
Campos 2017	88	2	23	2	0.79 [0.71, 0.86]	0.50 [0.07, 0.93]		
Carvalho 2018	27	5	2	6	0.93 [0.77, 0.99]	0.55 [0.23, 0.83]		
Castilho 2016	24	4	11	4	0.69 [0.51, 0.83]	0.50 [0.16, 0.84]		
Chawla 2012	44	4	38	49	0.54 [0.42, 0.65]	0.92 [0.82, 0.98]		
Cinar 2013	28	2	5	9	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]		
Costa 2002	6	0	26	17	0.19 [0.07, 0.36]	1.00 [0.80, 1.00]		
Da Silva 2010	28	5	0	17	1.00 [0.88, 1.00]	0.77 [0.55, 0.92]		
Diriz 2009	146	3	10	4	0.94 [0.89, 0.97]	0.57 [0.18, 0.90]		
Diriz 2011	77	1	21	5	0.79 [0.69, 0.86]	0.83 [0.36, 1.00]		
Diriz 2012	93	2	7	3	0.93 [0.86, 0.97]	0.60 [0.15, 0.95]		
Diriz 2019	57	7	9	15	0.86 [0.76, 0.94]	0.68 [0.45, 0.86]		
Ekstrand 2007	99	3	2	35	0.98 [0.93, 1.00]	0.92 [0.79, 0.98]		
Ekstrand 2011	92	8	0	12	1.00 [0.96, 1.00]	0.60 [0.36, 0.81]		
Ekstrand 2011a	118	17	0	16	1.00 [0.97, 1.00]	0.48 [0.31, 0.66]		
Erten 2005	14	4	39	27	0.26 [0.15, 0.40]	0.87 [0.70, 0.96]		
Forgie 2003	30	1	39	10	0.43 [0.32, 0.56]	0.91 [0.59, 1.00]		
Freitas 2016	62	14	51	39	0.55 [0.45, 0.64]	0.74 [0.60, 0.85]		
Goel 2009	39	0	42	2	0.48 [0.37, 0.60]	1.00 [0.16, 1.00]		
Haak 2002	199	34	14	73	0.93 [0.89, 0.96]	0.68 [0.59, 0.77]		
Hintze 2003	145	46	20	162	0.88 [0.82, 0.92]	0.78 [0.72, 0.83]		
Hintze 2003a	84	0	59	15	0.59 [0.50, 0.67]	1.00 [0.78, 1.00]		
Huth 2010	54	11	2	29	0.96 [0.88, 1.00]	0.72 [0.56, 0.85]		
Iranzo-Cortes 2017	40	0	9	15	0.82 [0.68, 0.91]	1.00 [0.78, 1.00]		
Iranzo-Cortes 2018	38	1	6	20	0.86 [0.73, 0.95]	0.95 [0.76, 1.00]		
Jablonski-Momeni 2008	65	7	6	15	0.92 [0.83, 0.97]	0.68 [0.45, 0.86]		
Jablonski-Momeni 2012	71	0	0	13	1.00 [0.95, 1.00]	1.00 [0.75, 1.00]		
Kim 2017	50	23	122	87	0.29 [0.22, 0.36]	0.79 [0.70, 0.86]		
Ko 2015	61	3	15	13	0.80 [0.70, 0.89]	0.81 [0.54, 0.96]		
Kockanat 2017	91	1	3	25	0.97 [0.91, 0.99]	0.96 [0.80, 1.00]		
Kucukyilmaz 2015	136	1	28	35	0.83 [0.76, 0.88]	0.97 [0.85, 1.00]		
Kuhnisch 2009	46	5	2	1	0.96 [0.86, 0.99]	0.17 [0.00, 0.64]		
Kuhnisch 2009a	53	5	0	7	1.00 [0.93, 1.00]	0.58 [0.28, 0.85]		
Laitala 2017	83	93	332	1595	0.20 [0.16, 0.24]	0.94 [0.93, 0.96]		
Mendes 2006	72	13	11	14	0.87 [0.78, 0.93]	0.52 [0.32, 0.71]		
Mitropoulos 2010	23	8	1	8	0.96 [0.79, 1.00]	0.50 [0.25, 0.75]		
Nakagawa 2013	91	7	4	25	0.96 [0.90, 0.99]	0.78 [0.60, 0.91]		
Nakajima 2014	20	1	8	9	0.71 [0.51, 0.87]	0.90 [0.55, 1.00]		
Neuhaus 2011	21	4	6	6	0.78 [0.58, 0.91]	0.60 [0.26, 0.88]		
Neuhaus 2015	63	5	22	8	0.74 [0.63, 0.83]	0.62 [0.32, 0.86]		
Novaes 2009	53	19	203	346	0.21 [0.16, 0.26]	0.95 [0.92, 0.97]		
Novaes 2010	314	31	164	83	0.66 [0.61, 0.70]	0.73 [0.64, 0.81]		
Novaes 2012	62	40	2	9	0.97 [0.89, 1.00]	0.18 [0.09, 0.32]		
Paula 2011	35	0	21	8	0.63 [0.49, 0.75]	1.00 [0.63, 1.00]		
Pereira 2011	45	11	10	30	0.82 [0.69, 0.91]	0.73 [0.57, 0.86]		
Qudeimat 2019	220	14	22	14	0.91 [0.87, 0.94]	0.50 [0.31, 0.69]		
Rocha 2003	24	3	5	18	0.83 [0.64, 0.94]	0.86 [0.64, 0.97]		
Rodrigues 2008	105	0	6	8	0.95 [0.89, 0.98]	1.00 [0.63, 1.00]		
Rodrigues 2009	129	10	7	2	0.95 [0.90, 0.98]	0.17 [0.02, 0.48]		
Rodrigues 2009a	162	7	5	5	0.97 [0.93, 0.99]	0.42 [0.15, 0.72]		
Rodrigues 2013	109	9	32	20	0.77 [0.69, 0.84]	0.69 [0.49, 0.85]		
Seremidi 2012	38	5	45	19	0.46 [0.35, 0.57]	0.79 [0.58, 0.93]		
Shi 2000	61	1	2	10	0.97 [0.89, 1.00]	0.91 [0.59, 1.00]		
Shimada 2010	76	5	19	11	0.80 [0.71, 0.88]	0.69 [0.41, 0.89]		
Sidi 1988	45	255	240	8558	0.16 [0.12, 0.21]	0.97 [0.97, 0.97]		
Souza 2013	45	5	15	14	0.75 [0.62, 0.85]	0.74 [0.49, 0.91]		

**Test 1. (Continued)**

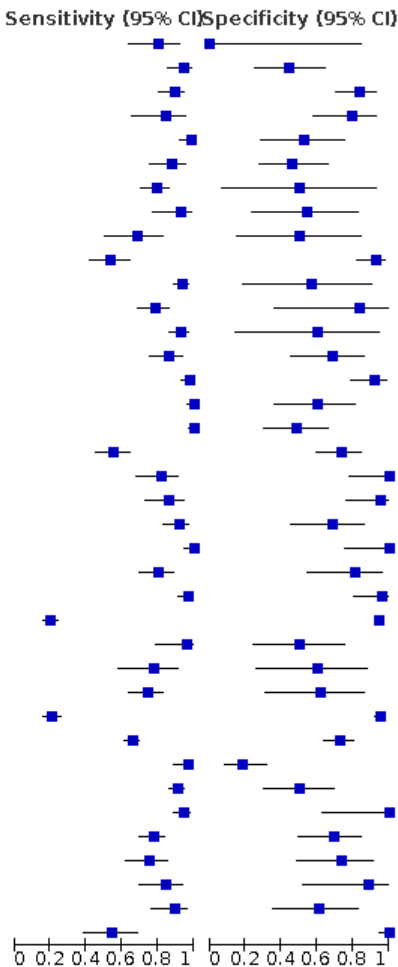
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Srinivasa 2010	70	0	19	11	0.80 [0.71, 0.89]	0.69 [0.41, 0.89]
Sidi 1988	45	255	240	8558	0.16 [0.12, 0.21]	0.97 [0.97, 0.97]
Souza 2013	45	5	15	14	0.75 [0.62, 0.85]	0.74 [0.49, 0.91]
Soviero 2012	33	1	6	8	0.85 [0.69, 0.94]	0.89 [0.52, 1.00]
Sridhar 2009	43	5	1	1	0.98 [0.88, 1.00]	0.17 [0.00, 0.64]
Teo 2014	41	7	5	11	0.89 [0.76, 0.96]	0.61 [0.36, 0.83]
Tonkaboni 2019	25	0	21	62	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]
Xiao-Hua 2016	57	10	15	15	0.79 [0.68, 0.88]	0.60 [0.39, 0.79]



**Test 2. ICDAS**

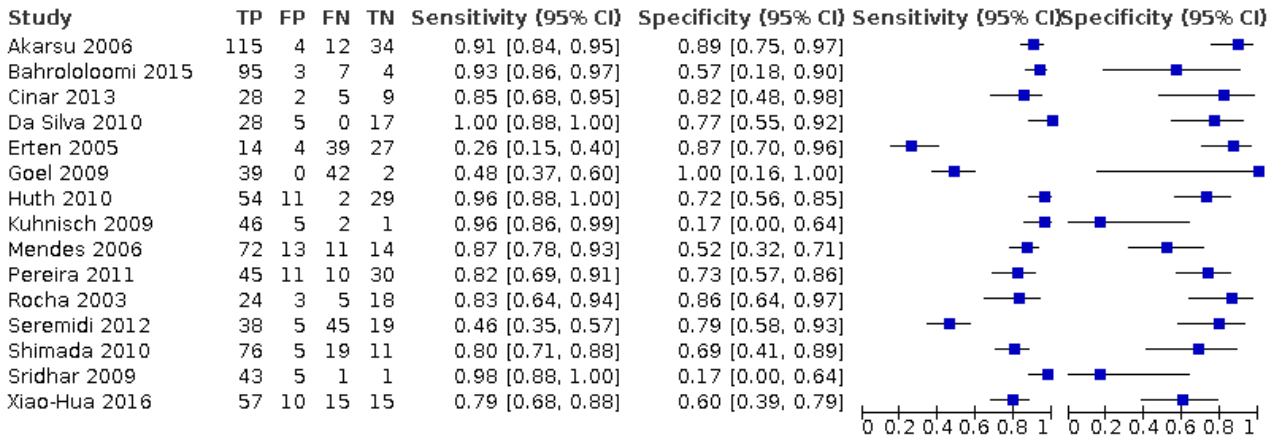
ICDAS

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	29	2	7	0	0.81 [0.64, 0.92]	0.00 [0.00, 0.84]
Bottenberg 2016	65	15	4	12	0.94 [0.86, 0.98]	0.44 [0.25, 0.65]
Braga 2009	73	8	9	41	0.89 [0.80, 0.95]	0.84 [0.70, 0.93]
Braga 2010	22	5	4	19	0.85 [0.65, 0.96]	0.79 [0.58, 0.93]
Braun 2017	64	9	1	10	0.98 [0.92, 1.00]	0.53 [0.29, 0.76]
Bussaneli 2015a	43	15	6	13	0.88 [0.75, 0.95]	0.46 [0.28, 0.66]
Campos 2017	88	2	23	2	0.79 [0.71, 0.86]	0.50 [0.07, 0.93]
Carvalho 2018	27	5	2	6	0.93 [0.77, 0.99]	0.55 [0.23, 0.83]
Castilho 2016	24	4	11	4	0.69 [0.51, 0.83]	0.50 [0.16, 0.84]
Chawla 2012	44	4	38	49	0.54 [0.42, 0.65]	0.92 [0.82, 0.98]
Diniz 2009	146	3	10	4	0.94 [0.89, 0.97]	0.57 [0.18, 0.90]
Diniz 2011	77	1	21	5	0.79 [0.69, 0.86]	0.83 [0.36, 1.00]
Diniz 2012	93	2	7	3	0.93 [0.86, 0.97]	0.60 [0.15, 0.95]
Diniz 2019	57	7	9	15	0.86 [0.76, 0.94]	0.68 [0.45, 0.86]
Ekstrand 2007	99	3	2	35	0.98 [0.93, 1.00]	0.92 [0.79, 0.98]
Ekstrand 2011	92	8	0	12	1.00 [0.96, 1.00]	0.60 [0.36, 0.81]
Ekstrand 2011a	118	17	0	16	1.00 [0.97, 1.00]	0.48 [0.31, 0.66]
Freitas 2016	62	14	51	39	0.55 [0.45, 0.64]	0.74 [0.60, 0.85]
Iranzo-Cortes 2017	40	0	9	15	0.82 [0.68, 0.91]	1.00 [0.78, 1.00]
Iranzo-Cortes 2018	38	1	6	20	0.86 [0.73, 0.95]	0.95 [0.76, 1.00]
Jablonski-Momeni 2008	65	7	6	15	0.92 [0.83, 0.97]	0.68 [0.45, 0.86]
Jablonski-Momeni 2012	71	0	0	13	1.00 [0.95, 1.00]	1.00 [0.75, 1.00]
Ko 2015	61	3	15	13	0.80 [0.70, 0.89]	0.81 [0.54, 0.96]
Kockanat 2017	91	1	3	25	0.97 [0.91, 0.99]	0.96 [0.80, 1.00]
Laitala 2017	83	93	332	1595	0.20 [0.16, 0.24]	0.94 [0.93, 0.96]
Mitropoulos 2010	23	8	1	8	0.96 [0.79, 1.00]	0.50 [0.25, 0.75]
Neuhaus 2011	21	4	6	6	0.78 [0.58, 0.91]	0.60 [0.26, 0.88]
Neuhaus 2015	63	5	22	8	0.74 [0.63, 0.83]	0.62 [0.32, 0.86]
Novaes 2009	53	19	203	346	0.21 [0.16, 0.26]	0.95 [0.92, 0.97]
Novaes 2010	314	31	164	83	0.66 [0.61, 0.70]	0.73 [0.64, 0.81]
Novaes 2012	62	40	2	9	0.97 [0.89, 1.00]	0.18 [0.09, 0.32]
Qudeimat 2019	220	14	22	14	0.91 [0.87, 0.94]	0.50 [0.31, 0.69]
Rodrigues 2008	105	0	6	8	0.95 [0.89, 0.98]	1.00 [0.63, 1.00]
Rodrigues 2013	109	9	32	20	0.77 [0.69, 0.84]	0.69 [0.49, 0.85]
Souza 2013	45	5	15	14	0.75 [0.62, 0.85]	0.74 [0.49, 0.91]
Soviero 2012	33	1	6	8	0.85 [0.69, 0.94]	0.89 [0.52, 1.00]
Teo 2014	41	7	5	11	0.89 [0.76, 0.96]	0.61 [0.36, 0.83]
Tonkaboni 2019	25	0	21	62	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]



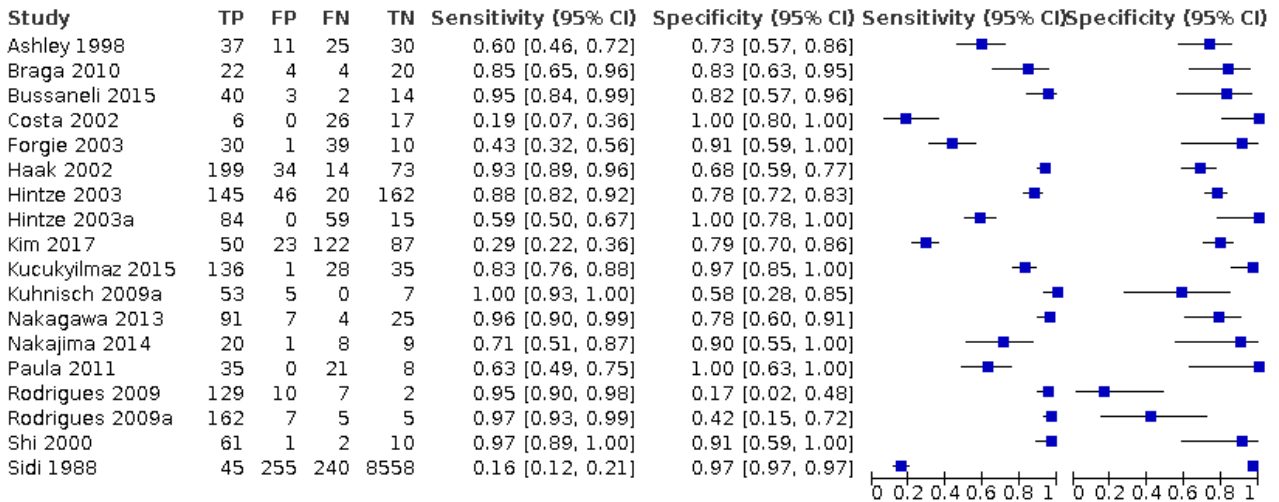
**Test 3. ERK**

ERK



**Test 4. Other visual**

Other visual



**ADDITIONAL TABLES**

**Table 1. Classification of levels of caries levels**

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

**Table 1. Classification of levels of caries levels** (Continued)

D <sub>4</sub>	Cavity progressing into pulp
----------------	------------------------------

DMFT = decayed, missing, and filled teeth.

**Table 2. Common visual scoring systems**

Name of scoring system	Description	Scale
<b>ICDAS</b>	Classifies the severity and activity of lesions on primary and permanent teeth; occlusal, approximal or smooth; and those adjacent to restorations. The surface is the unit of interest. Examination uses a dental mirror, ball-ended probe and 3-in-1 syringe to dry surfaces. There are 2 digits, the surface condition and the caries state (we will focus on the caries state numbering here). Codes also suggest treatment of management methods. Codes 1-2 the white/brown spot lesions (non-cavitated) require non-invasive treatment. Codes 3-4 the shadowed enamel cavitation may require restorations. Codes 5-6 the dentine cavitation may require more severe treatment or extraction ( <a href="#">Ismail 2007</a> ; <a href="#">Pitts 2001</a> )	0: sound 1: first visual change in enamel 2: distinct visual change in enamel 3: localised enamel breakdown 4: underlying dark shadow from dentine 5: distinct cavity with visible dentine 6: extensive distinct cavity with visible dentine
<b>ERK</b>	Aimed to detect and classify 'hidden caries' that are visible on bitewing radiographs but often has no obvious visual signs of caries. The scale relies on the classification of enamel translucency before and after drying, in addition to opacity, discolouration which may be indicative of underlying caries in the dentine ( <a href="#">Ekstrand 1997</a> )	0: no or slight changes in enamel translucency (after drying)  1: opacity or discolouration hardly visible on wet surface but distinctly visible when dry  2: opacity or discolouration visible without drying  3: localised enamel breakdown  4: cavitation exposing dentine
<b>Nyvad</b>	Based on surface characteristics rather than lesion depth to identify 2 levels of severity - non-cavitated and cavitated, the former focusing on enamel lesions. Examination should be performed on teeth free of dental plaque, surfaces should be dried for 10 to 15 seconds. A sharp explorer is generally used to assess the texture and roughness of the lesion; translucency/opacity, location, and colour can be assessed visually. Can be used on occlusal, approximal and smooth surfaces, also includes scores for secondary and root caries ( <a href="#">Zandona 2019</a> )	0 = sound: normal enamel translucency and texture  1 = active non-cavitated: whitish, or yellowish appearance; feel rough when explorer is gently moved across the surface  2 = active cavitated (surface discontinuity): localised surface defect (microcavity) in enamel only  3 = active cavitated (surface discontinuity): present similar surface characteristics but will feel softer rather than rougher, cavity easily observed  4 = inactive cavitated (intact surface): whitish, brownish or black; shiny, feels hard and smooth  5 = inactive caries (surface discontinuity): localised surface defect (microcavity) in enamel only

**Table 2. Common visual scoring systems** (Continued)

6 = inactive caries (cavity): easily visible cavity, surface black, shiny and hard

7 = filling (sound surface)

8 = filling + active caries non-cavitated

9 = filling + inactive caries (Nyvad 1999)

ERK = Ekstrand-Ricketts-Kidd system; ICDAS = International Caries Detection and Assessment System.

**Table 3. QUADAS-2 tool**

Item	Response (delete as required)
<b>Participant selection – Risk of bias</b>	
1) Was a consecutive or random sample of participants or teeth used?	<p><b>Yes</b> – where teeth or participants were selected consecutively or allocated to the study via a randomisation process</p> <p><b>No</b> – if study described another method of sampling</p> <p><b>Unclear</b> – if participant sampling is not described</p>
2) Was a case-control design avoided?	<p><b>Yes</b> – if case-control clearly not used</p> <p><b>No</b> – if study described as case-control or describes sampling specific numbers of participants with particular diagnoses</p> <p><b>Unclear</b> – if not clearly described</p>
3) Did the study avoid inappropriate exclusions (e.g. inclusion of caries into dentine)?	<p><b>Yes</b> – if the study clearly reports that included participants or teeth were apparently healthy or caries into dentine were excluded</p> <p><b>No</b> – if lesions were included that showed caries into dentine or exclusions that might affect test accuracy (e.g. teeth with no caries)</p> <p><b>Unclear</b> – if not clearly reported</p>
<b>Could the selection of participants have introduced bias?</b>	
If answers to all of questions 1) and 2) and 3) was 'yes'	<b>Risk is Low</b>
If answers to any of questions 1) or 2) or 3) was 'no'	<b>Risk is High</b>
Studies moved from low to unclear if answers to any of questions 1) and 2) and 3) was 'unclear'	<b>Risk is Unclear</b>
<b>Participant selection – Concerns regarding applicability</b>	
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><b>Yes</b> – if a group of participants or teeth has been included which is apparently healthy or indicative of early caries</p> <p><b>No</b> – if a group of participants or teeth has been included which is suspected of advanced caries</p>

**Table 3. QUADAS-2 tool** (Continued)

	<b>Unclear</b> – if insufficient details are provided to determine the spectrum of participants or teeth
2) Did the study report data on a per-patient rather than on a tooth or surface basis?	<b>Yes</b> – if the analysis was reported on a surface or tooth basis <b>No</b> – if the analysis was reported on a per-patient basis <b>Unclear</b> – 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?	<b>Yes</b> – if the participants were recruited prior to tooth extraction <b>No</b> – if previously extracted teeth were used in the analysis <b>Unclear</b> – if it was not possible to assess the source and method of recruiting of included participants/teeth
<b>Is there concern that the included participants or teeth do not match the review question?</b>	
If answers to all of questions 1) and 2) and 3) was 'yes'	<b>Risk is Low</b>
If answers to any of questions 1) or 2) or 3) was 'no'	<b>Risk is High</b>
Studies moved from low to unclear if answers to any of questions 1) and 2) and 3) was 'unclear'	<b>Risk is Unclear</b>
<b>Index test - Risk of bias (to be completed per test evaluated)</b>	
1) Was the index test result interpreted without knowledge of the results of the reference standard?	<b>Yes</b> – 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 <b>No</b> – if index test described as interpreted in knowledge of reference standard result <b>Unclear</b> – if index test blinding is not described
2) Was the diagnostic threshold at which the test was considered positive pre-specified?	<b>Yes</b> – if threshold was pre-specified (i.e. prior to analysing the study results) <b>No</b> – if threshold was not pre-specified <b>Unclear</b> – if not possible to tell whether or not diagnostic threshold was pre-specified
<b>For visual and radiograph tests only:</b> 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?	<b>Yes</b> – 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) <b>No</b> – if study states reported by same reader <b>Unclear</b> – if no mention of number of readers for each threshold or if pre-specification of threshold not reported <b>N/A</b> – multiple diagnostic thresholds not reported for the same index test
<b>Could the conduct or interpretation of the index test have introduced bias?</b>	
<b>For visual and radiographic studies item 3) to be added</b>	

**Table 3. QUADAS-2 tool** (Continued)

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



**Table 3. QUADAS-2 tool** (Continued)

Studies moved from low to unclear if the answer to question 1) or 2) was 'unclear'	<b>Concern is Unclear</b>
<b>Reference standard - Concerns regarding applicability</b>	
1) Does the study use the same definition of disease positive as the prescribed in the review question?	<p><b>Yes</b> - same definition of disease positive used, or teeth can be disaggregated and regrouped according to review definition</p> <p><b>No</b> - some teeth cannot be disaggregated</p> <p><b>Unclear</b> - definition of disease positive not clearly reported</p>
<b>Flow and timing - Risk of bias</b>	
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><b>Yes</b> - if study reports index and reference standard had a suitable interval or storage method</p> <p><b>No</b> - 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><b>Unclear</b> - 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><b>Yes</b> - if all participants underwent the same reference standard</p> <p><b>No</b> - if more than 1 reference standard was used</p> <p><b>Unclear</b> - if not clearly reported</p>
3) Were all participants included in the analysis?	<p><b>Yes</b> - if all participants were included in the analysis</p> <p><b>No</b> - if some participants were excluded from the analysis</p> <p><b>Unclear</b> - if not clearly reported</p>
If answers to questions 1) and 2) and 3) was 'yes'	<b>Risk is Low</b>
If answers to any one of questions 1) or 2) or 3) was 'no'	<b>Risk is High</b>
Studies moved from low to unclear if answers to any of questions 1) or 2) or 3) was 'unclear'	<b>Risk is Unclear</b>

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

**Table 4. Study characteristics**

Test	Study ID	Prevalence of enamel caries	Prevalence of caries in-to dentine	Dentition	Reference standard	Surface	Setting	Multisite
ICDAS	Achilleos 2013	0.95	High	Permanent	Histology	Occlusal	In vitro	No
ERK	Akarsu 2006	0.77	High	Permanent	Excavation	Occlusal	Dental hospital	No
Other visual	Ashley 1998	0.6	High	Permanent	Histology	Occlusal	In vitro	No
ERK	Bahrololoomi 2015	0.94	High	Primary	Excavation	Occlusal	Dental hospital	No
ICDAS	Bottenberg 2016	0.75	High	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Braga 2009	0.63	Medium	Primary	Histology	Approximal	In vitro	Yes
ICDAS	Braga 2010	0.52	Medium	Primary	Histology	Occlusal	In vivo - non-clinic setting	No
Other visual	Braga 2010 2 test	0.52	Medium	Primary	Histology	Occlusal	In vivo - non-clinic setting	No
ICDAS	Braun 2017	0.77	High	Permanent	Histology	Occlusal	In vitro	No
Other visual	Bussaneli 2015	0.71	High	Primary	Separation	Approximal	Dental hospital	No
ICDAS	Bussaneli 2015a	0.63	Low	Primary	Histology	Occlusal	In vitro	No
ICDAS	Campos 2017	0.96	High	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Carvalho 2018	0.73	High	Permanent	Radiograph	Occlusal	In vitro	No
ICDAS	Castilho 2016	0.81	Low	Permanent	Histology	Occlusal	Clinic	No
ICDAS	Chawla 2012	0.61	Medium	Primary	Histology	Approximal	In vitro	No
ERK	Cinar 2013	0.75	Medium	Primary	Histology	Occlusal	Dental hospital	No
Other visual	Costa 2002	0.65	Medium	Permanent	Histology	Occlusal	In vitro	No
ERK	Da Silva 2010	0.56	High	Primary	Histology	Occlusal	In vitro	No

**Table 4. Study characteristics** *(Continued)*

ICDAS	Diniz 2009	0.96	High	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Diniz 2011	0.94	Medium	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Diniz 2012	0.95	Medium	Permanent	Histology	Occlusal	Dental hospital	No
ICDAS	Diniz 2019	0.75	High	Primary	Histology	Occlusal	In vitro	No
ICDAS	Ekstrand 2007	0.72	High	Primary	Histology	Occlusal	In vitro	No
ICDAS	Ekstrand 2011	0.78	High	Primary	Histology	Approximal	In vitro	Yes
ICDAS	Ekstrand 2011a	0.78	Medium	Permanent	Histology	Approximal	In vitro	Yes
ERK	Erten 2005	0.63	Medium	Permanent	Histology	Occlusal	In vitro	No
Other visual	Forgie 2003	0.86	High	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Freitas 2016	0.68	Medium	Primary	Histology	Approximal	Clinic	Yes
ERK	Goel 2009	0.54	High	Primary	Histology	Occlusal	Clinic	No
Other visual	Haak 2002	0.67	High	Permanent	Histology	Approximal	In vitro	Yes
Other visual	Hintze 2003	0.26	Low	Permanent	Histology	Approximal	In vitro	Yes
Other visual	Hintze 2003a	0.91	High	Permanent	Histology	Occlusal	In vitro	No
ERK	Huth 2010	0.66	High	Permanent	Excavation	Approximal	Clinic	No
ICDAS	Iranzo-Cortes 2017	0.77	Medium	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Iranzo-Cortes 2018	0.68	Not reported	Primary	Histology	Occlusal	In vitro	No
ICDAS	Jablonski-Momeni 2008	0.76	High	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Jablonski-Momeni 2012	0.84	High	Permanent	Excavation	Occlusal	In vitro	No
Other visual	Kim 2017	0.61	Medium	Permanent	Radiograph	Approximal	Clinic	No
ICDAS	Ko 2015	0.8	Medium	Permanent	Histology	Approximal	In vitro	No

**Table 4. Study characteristics** *(Continued)*

ICDAS	Kockanat 2017	0.78	Medium	Primary	Histology	Occlusal	Clinic	No
Other visual	Kucukyilmaz 2015	0.82	Medium	Primary	Histology	Occlusal	Clinic	No
ERK	Kuhnisch 2009	0.89	High	Permanent	Histology	Occlusal	In vitro	No
Other visual	Kuhnisch 2009a	0.82	High	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Laitala 2017	0.2	Low	Permanent	Radiograph	Approximal	Clinic	Yes
ERK	Mendes 2006	0.75	Medium	Primary	Histology	Occlusal	In vitro	Yes
ICDAS	Mitropoulos 2010	0.6	High	Permanent	Histology	Approximal	In vitro	Yes
Other visual	Nakagawa 2013	0.75	Medium	Primary	Histology	Smooth	In vitro	Yes
Other visual	Nakajima 2014	0.74	Medium	Primary	Histology	Occlusal	In vitro	Yes
ICDAS	Neuhaus 2011	0.73	Medium	Primary	Histology	Occlusal	In vitro	No
ICDAS	Neuhaus 2015	0.87	Medium	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Novaes 2009	0.41	Low	Primary	Separation	Approximal	Clinic	Yes
ICDAS	Novaes 2010	0.81	Low	Primary	Separation	Approximal	Clinic	Yes
ICDAS	Novaes 2012	0.57	Medium	Primary	Histology	Occlusal	In vitro	Yes
Other visual	Paula 2011	0.88	Medium	Permanent	Histology	Occlusal	In vitro	Yes
ERK	Pereira 2011	0.57	Medium	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Qudeimat 2019	0.9	Medium	Permanent	Histology	Occlusal	In vitro	No
ERK	Rocha 2003	0.58	Low	Primary	Histology	Occlusal	Clinic	No
ICDAS	Rodrigues 2008	0.93	High	Permanent	Histology	Occlusal	In vitro	No
Other visual	Rodrigues 2009	0.92	Low	Primary	Histology	Occlusal	In vitro	No
Other visual	Rodrigues 2009a	0.93	Low	Permanent	Histology	Occlusal	In vitro	No



**Table 4. Study characteristics** (Continued)

ICDAS	Rodrigues 2013	0.83	High	Permanent	Histology	Occlusal	In vitro	No
ERK	Seremidi 2012	0.78	Medium	Permanent	Histology	Occlusal	In vitro	Yes
Other visual	Shi 2000	0.73	High	Permanent	Histology	Smooth	In vitro	No
ERK	Shimada 2010	0.86	High	Permanent	Histology	Occlusal	In vitro	Yes
Other visual	Sidi 1988	0.03	Low	Permanent	Radiograph	Approximal	Clinic	Yes
ICDAS	Souza 2013	0.76	High	Primary	Histology	Occlusal	In vitro	No
ICDAS	Soviero 2012	0.81	Medium	Primary	Histology	Approximal	In vitro	Yes
ERK	Sridhar 2009	0.96	Low	Permanent	Histology	Occlusal	In vitro	No
ICDAS	Teo 2014	0.72	Medium	Primary	Histology	Occlusal	Dental hospital	No
ICDAS	Tonkaboni 2018	0.42	High	Permanent	Histology	Approximal	In vitro	No
ERK	Xiao-Hua 2016	0.74	High	Permanent	Histology	Occlusal	In vitro	No

ERK = Ekstrand-Ricketts-Kidd system; ICDAS = International Caries Detection and Assessment System.

## APPENDICES

### Appendix 1. MEDLINE Ovid search strategy

1. exp Tooth demineralization/
2. (teeth adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
3. (tooth adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
4. (dental adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
5. (enamel adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
6. (dentin adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
7. (root adj5 (cavit\$ or caries or carious or decay\$ or lesion\$ or deminerali\$ or reminerali\$)).mp.
8. Dental caries activity tests/
9. or/1-8
10. Physical examination/
11. ((visual or tactile) adj3 (exam\$ or inspect\$)).mp.
12. ((caries or "dental decay" or "tooth decay" or carious) adj3 (diagnos\$ or detect\$ or check\$ or assess\$)).mp.
13. ((diagnos\$ or detect\$) adj3 method).mp.
14. ("assessment system" or ICDAS or "Dundee Selectable Threshold" or "WHO criteria" or "World health organization criteria" or "Universal Visual Scoring System" or ERK).mp.
15. or/10-14
16. 9 and 15

### 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. physical examination/
10. ((visual or tactile) adj3 (exam\$ or inspect\$)).mp.
11. ((caries or "dental decay" or "tooth decay" or carious) adj3 (diagnos\$ or detect\$ or check\$ or assess\$)).mp.
12. ((diagnos\$ or detect\$) adj3 method).mp.
13. ("assessment system" or ICDAS or "Dundee Selectable Threshold" or "WHO criteria" or "World health organization criteria" or "Universal Visual Scoring System" or ERK).mp.
14. or/9-13
15. 8 and 14

### 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 ( visual OR tactile )

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

caries AND visual OR caries AND tactile

## WHAT'S NEW

Date	Event	Description
16 December 2021	Amended	Minor edit to external source of support

## HISTORY

Review first published: Issue 6, 2021

## CONTRIBUTIONS OF AUTHORS

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

Drafting the protocol: Richard Macey (RM) and Tanya Walsh (TW).

Developing the search strategy and screening of studies: TW, RM, Helen V Worthington (HVW), Anne-Marie Glenny (AMG), Philip Riley (PR), Lucy O'Malley (LM).

Data extraction and risk of bias assessment: TW, RM, HW, AMG, PR, LM.

Data analysis: TW, RM.

Interpretation and drafting of the final review: TW, RM, HW, AMG, PR, LM, David Ricketts (DR), Janet E Clarkson (JEC).

## DECLARATIONS OF INTEREST

Richard Macey: none known.

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

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

Anne-Marie Glenny: none known. I am Co-ordinating Editor of Cochrane Oral Health.

Helen V Worthington: none known. I am Statistical Editor with Cochrane Oral Health.

Lucy O'Malley: none known. I am Editor with 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

This project was funded by the National Institute for Health Research (NIHR Cochrane Programme Grant 16/114/23 'Detection and Diagnosis of Common Oral Diseases: Diagnostic Test Accuracy of Tests of Oral Cancer and Caries'). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care

- 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](http://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

## DIFFERENCES BETWEEN PROTOCOL AND REVIEW

For this review we have removed the objective stated in the generic protocol "in different roles (adjunct to visual examination or independent test)," as this is not relevant to the visual classification system review.

It was stated in the protocol that in vitro and in vivo studies would be compared, but since 49 of the 60 studies using histology were in vitro this analysis has not been completed. A comparison of the histology studies comparing the in vitro and in vivo studies has been performed as a sensitivity analysis to assess whether the additional in vivo studies make a meaningful difference to this group of studies.

## INDEX TERMS

### Medical Subject Headings (MeSH)

Bias; Confidence Intervals; Dental Caries [\*diagnosis]; \*Dental Enamel; \*Early Diagnosis; Palpation [\*methods]; Physical Examination [\*methods]; Sensitivity and Specificity

**MeSH check words**

Adult; Child; Humans