Cochrane Library

Cochrane Database of Systematic Reviews

Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions (Review)

Walsh T, Macey R, Kerr AR, Lingen MW, Ogden GR, Warnakulasuriya S

Walsh T, Macey R, Kerr AR, Lingen MW, Ogden GR, Warnakulasuriya S.
Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions.
Cochrane Database of Systematic Reviews 2021, Issue 7. Art. No.: CD010276.
DOI: 10.1002/14651858.CD010276.pub3.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions

Tanya Walsh[1], Richard Macey[1], Alexander R Kerr[2], Mark W Lingen[3], Graham R Ogden[4], Saman Warnakulasuriya[5]

[1]Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. [2]Department of Oral and Maxillofacial Pathology, Radiology and Medicine, New York University College of Dentistry, New York, USA. [3]Pritzker School of Medicine, Division of Biological Sciences, Department of Pathology, University of Chicago, Chicago, Illinois, USA. [4]Division of Oral and Maxillofacial Clinical Sciences, School of Dentistry, University of Dundee, Dundee, UK. [5]Clinical and Diagnostic Sciences, King's College London, London, UK

Contact: Tanya Walsh, 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: Walsh T, Macey R, Kerr AR, Lingen MW, Ogden GR, Warnakulasuriya S. Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions. *Cochrane Database of Systematic Reviews* 2021, Issue 7. Art. No.: CD010276. DOI: 10.1002/14651858.CD010276.pub3.

ABSTRACT

Background

Squamous cell carcinoma is the most common form of malignancy of the oral cavity, and is often proceeded by oral potentially malignant disorders (OPMD). Early detection of oral cavity squamous cell carcinoma (oral cancer) can improve survival rates. The current diagnostic standard of surgical biopsy with histology is painful for patients and involves a delay in order to process the tissue and render a histological diagnosis; other diagnostic tests are available that are less invasive and some are able to provide immediate results. This is an update of a Cochrane Review first published in 2015.

Objectives

Primary objective: to estimate the diagnostic accuracy of index tests for the detection of oral cancer and OPMD, in people presenting with clinically evident suspicious and innocuous lesions.

Secondary objective: to estimate the relative accuracy of the different index tests.

Search methods

Cochrane Oral Health's Information Specialist searched the following databases: MEDLINE Ovid (1946 to 20 October 2020), and Embase Ovid (1980 to 20 October 2020). The US National Institutes of Health Ongoing Trials Register (ClinicalTrials.gov) and the World Health Organization International Clinical Trials Registry Platform were also searched for ongoing trials to 20 October 2020. No restrictions were placed on the language or date of publication when searching the electronic databases. We conducted citation searches, and screened reference lists of included studies for additional references.

Selection criteria

We selected studies that reported the diagnostic test accuracy of the following index tests when used as an adjunct to conventional oral examination in detecting OPMD or oral cavity squamous cell carcinoma: vital staining (a dye to stain oral mucosa tissues), oral cytology, light-based detection and oral spectroscopy, blood or saliva analysis (which test for the presence of biomarkers in blood or saliva).

Data collection and analysis

Two review authors independently screened titles and abstracts for relevance. Eligibility, data extraction and quality assessment were carried out by at least two authors, independently and in duplicate. Studies were assessed for methodological quality using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2). Meta-analysis was used to combine the results of studies for each index test using the bivariate approach to estimate the expected values of sensitivity and specificity.

Main results

This update included 63 studies (79 datasets) published between 1980 and 2020 evaluating 7942 lesions for the quantitative meta-analysis. These studies evaluated the diagnostic accuracy of conventional oral examination with: vital staining (22 datasets), oral cytology (24 datasets), light-based detection or oral spectroscopy (24 datasets). Nine datasets assessed two combined index tests. There were no eligible diagnostic accuracy studies evaluating blood or salivary sample analysis. Two studies were classed as being at low risk of bias across all domains, and 33 studies were at low concern for applicability across the three domains, where patient selection, the index test, and the reference standard used were generalisable across the population attending secondary care.

The summary estimates obtained from the meta-analysis were:
- vital staining: sensitivity 0.86 (95% confidence interval (CI) 0.79 to 0.90) specificity 0.68 (95% CI 0.58 to 0.77), 20 studies, sensitivity low-certainty evidence, specificity very low-certainty evidence;
- oral cytology: sensitivity 0.90 (95% CI 0.82 to 0.94) specificity 0.94 (95% CI 0.88 to 0.97), 20 studies, sensitivity moderate-certainty evidence, specificity moderate-certainty evidence;
- light-based: sensitivity 0.87 (95% CI 0.78 to 0.93) specificity 0.50 (95% CI 0.32 to 0.68), 23 studies, sensitivity low-certainty evidence, specificity very low-certainty evidence; and
- combined tests: sensitivity 0.78 (95% CI 0.45 to 0.94) specificity 0.71 (95% CI 0.53 to 0.84), 9 studies, sensitivity very low-certainty evidence, specificity very low-certainty evidence.

Authors' conclusions

At present none of the adjunctive tests can be recommended as a replacement for the currently used standard of a surgical biopsy and histological assessment. Given the relatively high values of the summary estimates of sensitivity and specificity for oral cytology, this would appear to offer the most potential. Combined adjunctive tests involving cytology warrant further investigation. Potentially eligible studies of blood and salivary biomarkers were excluded from the review as they were of a case-control design and therefore ineligible. In the absence of substantial improvement in the tests evaluated in this updated review, further research into biomarkers may be warranted.

PLAIN LANGUAGE SUMMARY

What are the most accurate tests for finding cancer of the mouth (oral cancer) and conditions that may lead to oral cancer?

Why is it important to improve the detection of oral cancer?

Persistent abnormal patches or sores in the mouth may represent mouth cancer or oral potentially malignant disorders (OPMDs). OPMDs can sometimes turn into mouth cancer, but if identified early enough patient outcomes can be improved.

What is the aim of this review?

Diagnosing mouth cancer involves the surgical removal of a piece of affected tissue (biopsy) that is then sent to a laboratory for examination of the cells using a microscope. This is painful for patients and involves a delay in finding out the results. The aim of this Cochrane Review review was to find out the accuracy of less invasive diagnostic tests that may provide more timely results. Researchers in Cochrane included 63 studies to answer this question.

What was studied in the review?

Three tests used in addition to a visual examination were evaluated.

- Vital stain: a liquid that can be used as a mouthrinse or applied directly to a suspected abnormal area of the mouth. It is thought that any area that is coloured after applying this liquid has a high chance of being mouth cancer or an OPMD.
- Oral cytology: a brush is used to remove cells from the suspected abnormal area that are sent to a laboratory for microscopic examination.
- Light-based detection: a special light shone in the mouth that is believed to make cancerous areas appear different to healthy areas.

A small number of studies evaluated a combination of these tests. No studies evaluated the accuracy of tests of blood or saliva.

What are the main results of the review?

The review included 63 studies involving 7942 abnormal patches or mouth sores. Each study participant underwent one or more diagnostic tests as well as a surgical biopsy.

The proportion of people in the included studies with mouth cancer or OPMDs identified through surgical biopsy ranged widely from 4% to 97%. Based on adults attending general dental practices in the UK, in a sample of 1000 lesions 25 would be mouth cancer or an OPMD, and 975 lesions would not be mouth cancer or an OPMD.

- Of the 1000 lesions that are tested with vital staining: 22 will be correctly identified as having mouth cancer or an OPMD (true positives), but three lesions that truly are mouth cancer or an OPMD will remain undetected; their 'negative' test results will be incorrect (false negatives). 663 lesions will be correctly identified as not having mouth cancer or an OPMD (true negatives), but 312 people will be incorrectly identified; their 'positive' test results will suggest they have mouth cancer or an OPMD (false positives).

- For oral cytology: 23 lesions will be correctly identified as having mouth cancer or an OPMD (true positives), but two lesions that truly are mouth cancer or an OPMD will remain undetected (false negatives). 917 lesions will be correctly identified as not having mouth cancer or an OPMD (true negatives), but 58 lesions will be incorrectly identified (false positives).

- For light-based tests: 22 lesions will be correctly identified as having mouth cancer or an OPMD (true positives), but three lesions that truly are mouth cancer or an OPMD will remain undetected (false negatives). 488 lesions will be correctly identified as not having mouth cancer or an OPMD (true negatives), but 487 lesions will be incorrectly identified (false positives).

There was considerable variability in the accuracy of the vital staining and light-based tests which meant that the results of a future study could take a broad range of values.

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

There were shortcomings in many of the studies that put them at high risk of bias. We rated the certainty of the evidence as moderate for oral cytology and low or very low for the remaining tests.

Who do the results of this review apply to?

Studies included in the review were carried out in many different countries but no studies originated in Africa. Most studies were completed in hospitals. Studies were published between 1980 and 2020.

What are the implications of this review?

Although cytology was the most accurate of all the tests, none can be recommended as a replacement for the currently used standard of a surgical biopsy and pathology assessment. Most studies investigated patients that had been referred to a hospital clinic for further investigation and so we have only limited information on how accurate they would be when used by a general dentist or frontline medical provider.

How up-to-date is this review?

Review authors searched for and used studies published up to 20 October 2020.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings table - test comparisons

Question	What is the diagnostic accuracy of tests for the diagnosis of oral cancer and potentially malignant disorders?
Population	Patients referred to a secondary care facility for further investigation of a clinically evident lesion
Index test	Intra- and extra-oral conventional visual and tactile examination and adjunctive tests (vital staining, oral cytology, light-based examination, alone or in combination)
Comparator test	Estimates were compared across different adjuncts
Target condition	Oral cancer and potentially malignant disorders
Reference standard	Biopsy and histology
Action	If potentially malignant disorders can be correctly identified at an early stage then remedial action can be taken to improve morbidity and mortality
Diagnostic stage	Lesion detection or assessment of patients presenting with clinically evident suspicious or innocuous lesions
Quantity of evidence	63 studies providing 79 datasets. 4006 lesions with dysplasia or OSCC from a total of 7942 examined lesions (50% prevalence)

Findings: analysis comparing oral cytology, light-based examination, vital staining, combined tests (staining plus other adjunct) Consequences in a cohort of 1000 lesions

					Pre-test probability 2.5%[a]		Certainty of the evidence	
Test	**Datasets**	**Total lesions**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**False negatives (missed lesions)**	**False positives**	**Sensitivity**	**Specificity**
Oral cytology	24	2892	0.90 (0.82 to 0.94)	0.94 (0.88 to 0.97)	2 (1 to 4)	58 (29 to 117)	⊕⊕⊕⊖ MODERATE[b]	⊕⊕⊕⊖ MODERATE[b]
Light-based examination	24	2587	0.87 (0.78 to 0.93)	0.50 (0.32 to 0.68)	3 (2 to 5)	487 (312 to 663)	⊕⊕⊖⊖ LOW[c]	⊕⊖⊖⊖ VERY LOW[c]

Vital staining	22	1780	0.86 (0.79 to 0.90)	0.68 (0.58 to 0.77)	3 (2 to 5)	312 (224 to 409)	⊕⊕⊖⊖ LOW[d]	⊕⊖⊖⊖ VERY LOW[d]
Combined tests	9	683	0.78 (0.45 to 0.94)	0.71 (0.53 to 0.84)	5 (1 to 14)	283 (156 to 458)	⊕⊖⊖⊖ VERY LOW[e]	⊕⊖⊖⊖ VERY LOW[e]
Using vital staining as the reference category: difference in sensitivity for oral cytology -0.04 (-0.12 to 0.04) P = 0.33, light-based -0.02 (-0.11 to 0.08) P = 0.70, combined 0.08 (-0.18 to 0.33); difference in specificity for oral cytology -0.25 (-0.36 to -0.14) P < 0.001, light-based 0.18 (-0.03 to 0.39) P = 0.10, combined -0.03 (-0.21 to 0.16)								

[a]Hypothetical cohorts of 1000 lesions are presented for numbers estimated at a prevalence of 2.5%. The prevalence observed for any dysplasia reported in the included studies in the review is far higher than would be expected in the general population; accordingly we have used an illustrative prevalence of 2.5% based on evidence from patients presenting at general dental practices in the UK (Lim 2003; Warnakulasuriya 2011).

[b]We graded the evidence as moderate for sensitivity and specificity, downgrading once for study limitations.

[c]We graded the certainty of the evidence for sensitivity as low, downgrading 1 level for study limitations and 1 level for inconsistency; for specificity we graded the evidence as very low, additionally downgrading for imprecision.

[d]We graded the certainty of the evidence for sensitivity as low, downgrading 1 level for study limitations and 1 level for inconsistency; for specificity we graded the evidence as very low, additionally downgrading for imprecision.

[e]We graded the certainty of the evidence for sensitivity and specificity as very low, downgrading 1 level for study limitations, imprecision, and inconsistency.

CI: confidence interval; **OSCC:** oral cavity squamous cell carcinoma.

BACKGROUND

Target condition being diagnosed

The target conditions of interest are oral squamous cell carcinoma (OSCC) and oral potentially malignant disorders (OPMD) of the oral cavity. OSCC is the most common form of oral cavity cancer (Bagan 2020; Chi 2015) and a proportion is preceded by OPMD. OPMD represent a heterogeneous group of conditions including leukoplakia, erythroplakia, proliferative verrucous leukoplakia, lichen planus, oral submucous fibrosis, and actinic keratosis (Warnakulasuriya 2007).

The natural history of OSCC is not fully understood; not all OPMD undergo malignant transformation, some remain stable, and some affected sites can revert back to health (Speight 2017). Equally, some OSCC can develop from lesions in which epithelial dysplasia was not previously diagnosed (Dost 2014), or from apparently normal mucosa that may contain significant molecular aberrations that increase the likelihood of cancer (Farah 2019; Nikitakis 2018; Thompson 2017). Proliferative verrucous leukoplakia has the highest malignant transformation rate followed by erythroplakia (Locca 2020). Oral leukoplakia is the most common OPMD, but has a varied malignant transformation rate (MTR) (Arduino 2013; Warnakulasuriya 2016). Warnakulasuriya 2016 reported an MTR of oral leukoplakia to range between 0.13% and 34%. Chaturvedi 2019 and colleagues undertook the largest oral leukoplakia natural history study to date. They reported that the risk of progression to oral cancer increased with the grade of dysplasia, with a five-year competing risk-adjusted absolute risk for leukoplakia overall of 3.3% (95% confidence interval (CI) 2.7% to 3.9%); no dysplasia of 2.2% (95% CI 1.5% to 3.1%); mild dysplasia 11.9% (95% CI 7.1% to 18.1%); moderate dysplasia 8.7% (95% CI 3.2% to 17.9%); and severe dysplasia 32.2% (95% CI 8.1% to 60%). Petti 2003 calculated a global MTR of oral leukoplakia of 1.4% per year (95% CI 0.7% to 2%), but when this is applied to the prevalence of the condition, it far exceeds the numbers of actual cases of OSCC reported. However, the MTR in hospital-based studies is consistently higher than in community-based studies.

There is uncertainty regarding the risk of malignant transformation of oral lichen planus, with estimates between 0.1% to 1%, and this remains a controversial and unclear aspect of premalignant disease and oral cancer progression. Several recent systematic reviews have reported an MTR for oral lichen planus close to 1%. For example in a meta-analysis of 78 studies with 25,848 patients González-Moles 2019 reported a malignant transformation rate of 1.14% (95% CI 0.84% to 1.49%), results similar to Fitzpatrick 2014 (1.09%), Giuliani 2019 (1.37%), and Locca 2020 (1.4% (95% CI 0.9% to 1.9%)). In a meta-analysis of 33 studies with 12,838 oral lichen planus patients Idrees 2021 reported that 151 cases were initially considered to have progressed to carcinoma (1.2%). Following the application of strict criteria (the presence of a properly verified oral lichen planus diagnosis, a clear description of the cancerous lesion developing at the same site as the verified oral lichen planus lesion, and a follow-up period of a minimum of six months prior to carcinoma development), this figure was reduced to 0.44%, an MTR of 0.2% (95% CI 0.1% to 0.3%) (Idrees 2021).

The early detection and excision of high-risk oral leukoplakia can reduce the risk of malignant transformation (Mehanna 2009). Leukoplakias can be treated by a number of different methods although there remains some debate in the literature as to their effectiveness and there is limited empirical evidence (Holmstrup 2006; Lodi 2006). Systematic reviews have evaluated the evidence for surgical interventions (including laser therapy). Surgical laser excision of oral leukoplakia may decrease recurrence rates but have no effect on the malignant transformation of oral leukoplakia when compared with conventional treatments (de Pauli Paglioni 2020). There is little experimental evidence for non-surgical interventions and though clinical resolution was observed, relapses were common (Lodi 2016).

In the United Kingdom, patients presenting with any new growth, an ulcer, or a white and red, or red lesion, persisting for more than two to three weeks are urgently referred to Oral Medicine Units or Oral and Maxillofacial Surgery Units for further investigation (NICE 2016). Technologies to treat and manage oral cancer have progressed substantially (Bulsara 2018; Furness 2011; Glenny 2010; Shaw 2020). Despite this, mortality rates have remained high (approximately 50%) and have not improved in multiple European countries over several decades (Warnakulasuriya 2009; Warnakulasuriya 2020), and this appears to point to the aggressive biological behaviour of some oral cancers and a need for centralization of expertise whilst remaining accessible to the patient (Ogden 2020).

If the lesion is diagnosed as OSCC, the traditional treatment is surgery and radiotherapy, but the associated morbidity is high. This is in marked contrast to the improved survival rates in many other cancers, such as those of the breast and the colon (Cancer Research UK). Reasons for this include late presentation by the patient (early cancers are often asymptomatic, such that the patient is unaware of it) and delayed diagnosis (a combination of patient factors such as infrequent visits to their dentist or physician, and clinician factors, such as failure to screen the entire mouth, failure to raise their index of suspicion regarding any lesion they may see, or delays in onward referral) (Seoane 2016). Yet oral cancer can be cured if caught early enough (Ganly 2012).

Index test(s)

A number of index tests have been proposed as adjuncts to a conventional oral examination (COE) to improve diagnostic test accuracy (Fedele 2009; Kerr 2020; Leston 2010; Lingen 2008; Lingen 2017; Liu 2016; Madhura 2020; Omar 2015; Patton 2008; Rethman 2010).

- Vital staining (toluidine blue, tolonium chloride).
- Oral cytology (e.g. OralCDx brush biopsy).
- Light-based detection (e.g. ViziLite, Microlux/DL, VELscope, Orascoptic DK, Identafi 3000) and oral spectroscopy.
- Blood and saliva analysis.

We evaluated all four categories of index tests in this review (restricting vital staining to those applied to a visible lesion). All have the potential to be used as diagnostic or case-finding adjuncts to the COE by clinicians or other health professionals (Additional Table 1), to aid in the diagnosis of OSCC and OPMD.

Clinical pathway

There is no internationally recognised or standardised clinical pathway for individuals presenting with OPMD. Commonly, individuals receive a COE from frontline clinicians as part of a routine dental appointment. The COE involves a standard

visual and tactile examination of the oral mucosa under normal (incandescent) light. Alternatively, patients may occasionally present to a frontline clinician with symptoms. Upon discovering a lesion, the clinician makes a subjective judgement based upon the clinical presentation. If a malignancy or OPMD is suspected, the frontline clinician refers to an oral specialist for a surgical biopsy rendering a definitive diagnosis. In some healthcare systems, for example in Spain, the biopsy is performed in primary care by the dentist.

Supplementing the COE with an index test could aid the identification of clinically evident lesions. Tests could have a triage role in assisting the frontline clinician or oral specialist to more accurately assess oral lesions of uncertain significance. It could also help to reduce the unnecessary referral of benign conditions (Shi 2019). For instance, traumatic keratoses are common benign white lesions and referring every patient would be excessive and incur increased financial cost and patient anxiety. A non-invasive index test or combination of tests adjunctive to the COE could provide a frontline clinician with sufficient information to reduce the number of unnecessary referrals.

The index tests also have the potential to improve patient diagnosis in secondary care settings. Following referral to a specialist clinic, baseline surgical biopsy is commonly undertaken, and the specialist will take samples from areas that are clinically representative of the most severe disease. This becomes complicated when the lesion or lesions under investigation are large or heterogeneous in clinical appearance. Sample site selection may be facilitated by the use of diagnostic adjuncts. The tests could also be used to help monitor patients who have a history of oral cancer or OPMD and are at high risk for recurrence. This population has often been exposed to surgical procedures, repeated biopsies, and in some oral cancer cases, also to radiotherapy. Monitoring these patients for new and recurrent disease can become complicated and the diagnostic adjuncts could be of value.

Alternative test(s)

The COE has been investigated as a test for early detection of oral cancer and OPMD (Walsh 2013) and traditionally would be the first test applied. The only alternative tests are the adjunctive tests (which form the index tests of this review) and the surgical biopsy and histologic assessment which is the reference standard for this Cochrane Review.

Rationale

Oral cancer is a significant global health problem with an estimated 354,864 new cases and 177,384 deaths in 2018 (Bray 2018), and reported increases in incidence and mortality rates in many countries across the globe (Jin 2016; Shield 2017; Warnakulasuriya 2009). More recently, the trends of oral cancer incidence indicated two contrasting patterns between the sexes. In males, most cancer registry populations exhibited decreasing trends whilst in females, rising rates were seen in most populations (Miranda-Filho 2020). There is wide geographic variation in disease incidence and mortality, with almost double the incidence in low-income and middle-income countries compared to high-income countries, and a three-fold increase in mortality. Tobacco use, alcohol consumption, betel quid chewing, and low socio-economic status are the most important risk factors for oral cancer (IARC 2012;). Human papillomavirus is not considered a significant risk for oral cavity cancers but is a major risk factor for oropharynx cancer (Kreimer 2020). Men have a higher incidence of oral cancer than women, but the gender difference has narrowed in recent decades from a ratio of 5 males to 1 female diagnosed with oral cancers in the 1960s to less than 2 to 1 in 2008 (Ferlay 2010). Although traditionally the risk of oral cancer increases with age, since the 1980s the incidence amongst younger adults has increased in the European Union and the United States (Warnakulasuriya 2009).

Oral cancer mortality can be reduced by: (i) primary prevention, (ii) secondary prevention (screening and early detection), and (iii) improved treatment (ERO-FDI 2019). Accurate case detection and early treatment of oral cancers can substantially improve an individual's outlook with respect to morbidity, mortality, and quality of life (Speight 2017).

Unfortunately, oral cancer is often diagnosed at a late stage, when the prognosis is poor and the risks of significant morbidity and mortality are substantially higher (Rusthoven 2010). Technologies to treat and manage oral cancer have progressed substantially (Glenny 2010; Shaw 2020), but the five-year survival following diagnosis has remained constant at around 50% for the past 30 to 40 years (Bravi 2021; Cancer Research UK; Parkin 2001; Warnakulasuriya 2009).

In this review, we aimed to identify diagnostic tests for OSCC and OPMD and to evaluate the diagnostic accuracy of these tests (Additional Table 1) when used as adjuncts to a COE by clinicians in a primary or secondary care facility.

This diagnostic test accuracy review complements a number of intervention reviews undertaken by Cochrane Oral Health on the treatment of oral and oropharynx cancers (Bulsara 2018; Furness 2011; Glenny 2010) and oral leukoplakia (Lodi 2016), screening programmes for the early detection and prevention of oral cancer (Brocklehurst 2013), and clinical assessment to screen for the detection of oral cavity cancer and oral potentially malignant disorders in apparently healthy adults (Walsh 2013).

In this Cochrane Review we have included contemporary studies irrespective of publication language and status, and built upon existing systematic reviews to incorporate methodological developments by: expanding the search strategy to capture all relevant evidence, for tests developed and used globally; applying appropriate hierarchical analysis (Dinnes 2016); and assessing the body of evidence using GRADE (Schünemann 2020; Schünemann 2020a) to facilitate the production of summary of findings tables. The current version is an update of the original review first published in 2015 (Macey 2015).

OBJECTIVES

To estimate the diagnostic accuracy of index tests for the detection of oral cancer and oral potentially malignant disorders (OPMD) in people presenting with clinically evident suspicious and innocuous lesions.

Secondary objectives

To estimate the relative accuracy of the different index tests.

METHODS

Criteria for considering studies for this review

Types of studies

Studies evaluating index test(s) that reported on the diagnostic accuracy for individuals presenting with clinically evident suspicious or innocuous lesions. The index tests assessed in this review are provided in Additional Table 1. Eligible study designs included cross-sectional diagnostic test accuracy studies (or consecutive series) and randomised studies of diagnostic test accuracy. We excluded studies that collected data retrospectively, reported in abstract form alone, case-control studies, uncontrolled reports and randomised controlled trials (RCTs) of the effectiveness of screening programmes (intervention studies).

Participants

Adults (aged 16 years or over) presenting to primary or secondary care with clinically evident suspicious or innocuous oral lesions. The clinical setting is important as studies enrolling patients referred to a specialist care evaluate index tests that are predominantly carried out by clinical experts in cohorts with a higher percentage of dysplastic or malignant lesions than would be expected in a general population.

Index tests

Index tests used alone, or in combination, as an adjunct to the conventional oral examination (COE) (Additional Table 1) by either primary care clinicians or secondary care specialists. If a study reported two or more tests then we included those tests in the analysis if it was reasonable to assume that the two tests were conducted or interpreted independently of each other. If this was not the case, then we used the results from the first test only. Where an index test was used as a combined test and reported the results as a single data point, then these were considered as a combination of tests and reported and analysed separately.

Target conditions

Following the consensus views of the expert working group of the World Health Organization (WHO) Collaborating Centre for Oral Cancer/Precancer (Warnakulasuriya 2007), the following were considered.

- Oral cancer.
    - Oral cavity squamous cell carcinoma (OSCC).
- Oral potentially malignant disorders (OPMD).
    - Leukoplakia.
    - Proliferative verrucous leukoplakia.
    - Erythroplakia.
    - Lupus erythematosus.
    - Submucous fibrosis.
    - Hereditary disorders such as dyskeratosis congenita or epidermolysis bullosa.

This review classified any level of dysplasia (mild, moderate, or severe) or OSCC as the target conditions.

Reference standards

Surgical or punch biopsy with histological assessment. Studies not specifying verification by a pathology diagnosis as a reference standard were ineligible for inclusion.

Search methods for identification of studies

Electronic searches

Cochrane Oral Health's Information Specialist conducted systematic searches for diagnostic test accuracy studies in the following databases:

- MEDLINE Ovid (1946 to 20 October 2020) (Appendix 1);
- Embase Ovid (1980 to 20 October 2020) (Appendix 2).

The MEDLINE search strategy outlined in Appendix 1 was modified for the listed databases and was based on the companion Cochrane diagnostic test accuracy review (Walsh 2013). The search was not limited by language or publication status. Non-English studies were translated.

Searching other resources

The following trial registries were searched for ongoing studies:

- US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (clinicaltrials.gov; searched 20 October 2020) (Appendix 3);
- World Health Organization International Clinical Trials Registry Platform (apps.who.int/trialsearch; searched 20 October 2020) (Appendix 4).

We sought to locate further studies through citation searches and reference lists of key articles, and by contacting authors of identified articles to request information of any unpublished or ongoing studies. We identified the systematic review 'Adjuncts for the evaluation of potentially malignant disorders in the oral cavity' (Lingen 2017) published by the American Dental Association (ADA) through the collaboration of Cochrane Oral Health and the ADA.

For the previous version of this review, we searched the Cochrane Oral Health's Trials Register, Cochrane Diagnostic Test Accuracy Register, and the MEDION database on 30 April 2013. The latter two databases were no longer available to search for this update. The search strategies for these databases can be found in Appendix 5.

Data collection and analysis

Selection of studies

Two of the review authors independently assessed titles and abstracts of all identified studies from the electronic searches. Full reports were obtained for studies that appeared to meet the inclusion criteria, or where a clear decision could not be made from the title and abstract alone. Where disagreements occurred, we resolved these by discussion or by consulting a third review author.

Data extraction and management

Two review authors independently extracted data using a piloted data collection form and, where necessary, study authors were contacted to obtain relevant data missing from the full paper.

The following data were recorded from each study.

- Sample characteristics (age, sex, socio-economic status, risk factors where stated (e.g. human papillomavirus status positive/negative, prevalence of tobacco use and alcohol consumption), number of patients/lesions, lesion site).
- Setting (country, disease prevalence, type of facility).
- The type of index test(s) used (category, name, 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).

We had planned to extract data according to subgroup (tobacco and alcohol consumption) but these data were rarely reported in the studies.

Assessment of methodological quality

Two review authors independently assessed the quality of the included studies. Where disagreements occurred, these were either resolved by discussion or by consulting a third review author. We modified the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) (Whiting 2011), piloted it on five studies and then used it to assess the methodological quality of the diagnostic studies over four key domains: patient selection, index test, reference standard, and flow and timing of participants through the study (Additional Table 2).

Two core signalling questions were removed as they were addressed by the eligibility criteria: 'Was a case-control design avoided?' and 'Did all patients receive a reference standard?.' Three additional signalling items were added relating to commercial funding, training and calibration, and multiple index tests.

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

Results of the assessment of methodological quality were also presented graphically.

Statistical analysis and data synthesis

The threshold of interest was between a benign tissue diagnosis and any oral epithelial dysplasia or OSCC. 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 or lesions reported within a single study. When there were two or more test results reported in the same study, we included them as separate datasets, since the unit of analysis was the test result, not the patient. Diagnostic odds ratios (DOR), defined as the ratio of the odds of test positive in those who have the target condition relative to the odds of test positive in those without the condition, were calculated as a global measure of accuracy.

Hierarchical models were used for data synthesis. The data were extracted for the target condition of oral cancer and OPMD. 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 using bivariate models to estimate the summary values of sensitivity and specificity at a common threshold for each index test (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/meqrlogit and the METANDI package in Stata (Harbord 2009; Takwoingi 2016).

Investigations of heterogeneity

The investigation of each potential source of heterogeneity was considered individually for each index test. Initially, a visual inspection of the clinical and methodological characteristics of the included studies, coupled forest plots, and summary receiver operating characteristic (SROC) plots were used to form the basis of the assessment of heterogeneity. Where sufficient numbers of studies allowed, meta-regression analyses were conducted with xtmelogit/meqrlogit to explore possible sources of heterogeneity. We added the source of heterogeneity as covariates to the bivariate model, and used a likelihood ratio (LR) test to formally assess the significance of any model comparisons (Macaskill 2011; Takwoingi 2016). Initially we allowed the source of heterogeneity 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). We modelled different variances wherever possible.

The sources of heterogeneity (specified a priori) were prevalence of the disease in the study, and methods used by index test for example classification of fluorescence device for light-based studies. Each potential source of heterogeneity was investigated separately.

We had intended to investigate the severity of the target condition as another potential source of heterogeneity. As most studies included 'any dysplasia' as the target condition, there was insufficient variation in the included studies to do this.

Sensitivity analyses

We intended to carry out a sensitivity analyses that restricted the analysis to studies where the reference standard was surgical or punch biopsy followed by histopathology, however, all included studies used surgical biopsy as the reference standard.

Assessment of reporting bias

We did not undertake tests for reporting bias as these can be misleading when applied to systematic reviews of diagnostic test accuracy (Leeflang 2008; Tang 2000).

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 (Schünemann 2020; Schünemann 2020a), and using the GRADEPro online tool (www.guidelinedevelopment.org). To enhance readability and understanding, we presented test accuracy results in natural frequencies to indicate numbers of false positives and false negatives. The certainty of the body of evidence was assessed with

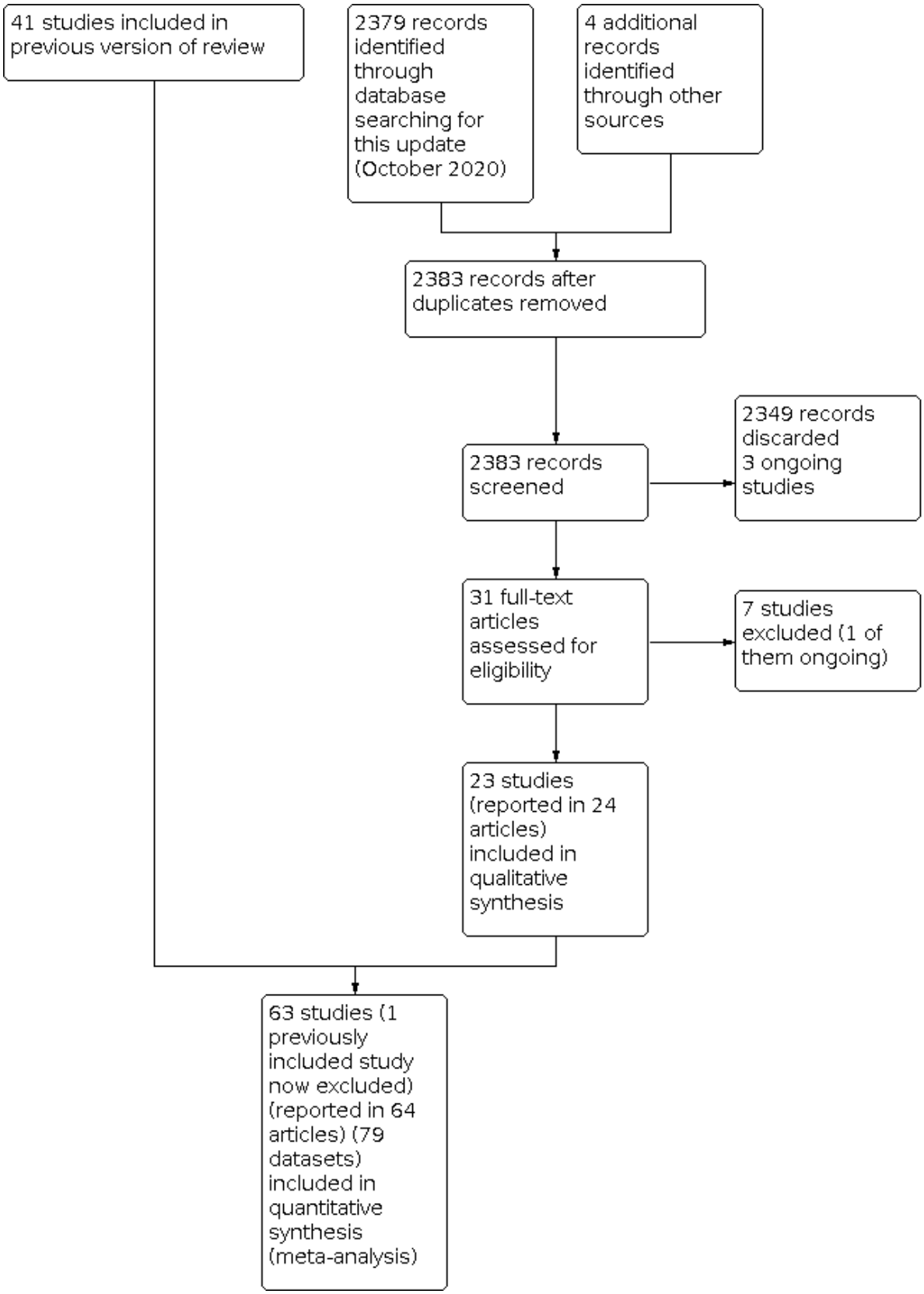
reference to the overall risk of bias of the included studies, the indirectness of the evidence, the inconsistency of the results, and the imprecision of the estimates. We categorised the certainty of the body of evidence, as high, moderate, low, or very low.

RESULTS

Results of the search

We retrieved 2379 references in the search update on October 2020, two records from the American Dental Association (ADA) systematic review, and two additional records through other sources. 2383 records remained after the removal of duplicates. These records were screened and 2349 records were discarded. Three studies are ongoing. 31 full-text articles were assessed for eligibility and seven were excluded. This update includes 23 new studies (24 articles) and 41 studies from the previous version, totaling 63 studies (64 articles) (one previously included study has been excluded) in 79 datasets for the quantitative meta-analysis (Figure 1).

Figure 1. Study flow diagram.



Included studies evaluated data from patients with a total of 7942 lesions. Most studies were conducted in India (17 studies), Germany (8 studies), and the US (7 studies). The remainder were conducted in Australia (3 studies), Brazil (2 studies), Canada (1 study), China (4 studies), Iran (2 studies), Italy (4 studies), Japan (2 studies), Poland (1 study), Spain (2 studies), Sri Lanka (3 studies), Taiwan (2 studies), Turkey (1 study), and the UK (4 studies). The studies were published between 1980 and 2020.

Of the four categories of index tests proposed for evaluation, we identified studies reporting the diagnostic test accuracy of vital staining, oral cytology, and light-based detection methods. No eligible studies assessed the diagnostic test accuracy of blood and saliva analysis in the detection of oral cavity squamous cell carcinoma (OSCC) or oral potentially malignant disorders (OPMD). All included studies were carried out in a secondary care (hospital) setting.

The majority of studies evaluated a single adjunct test on a single sample. Of the remainder, one assessed a single test on two different samples (Cheng 2003a; Cheng 2003b), one assessed two tests on two different samples (Mehrotra 2010), 10 assessed two tests on the same sample (Adil 2017; Awan 2011; Chainani-Wu 2015; Chaudhari 2013; Kaur 2016; Nanayakkara 2016; Petruzzi 2014; Rahman 2012; Shukla 2018; Ujaoney 2012), and two assessed three tests on the same sample (Kämmerer 2013; Yamamoto 2017). Nine studies evaluated index tests used in combination (Baeten 2018; Chainani-Wu 2015; Epstein 2008; Guneri 2011; Gupta 2007; Johnson 2019; Mehrotra 2010; Mojsa 2012; Ujaoney 2012).

For the purposes of evaluation.

- Twenty studies evaluated vital staining or rinsing. These studies provided 22 datasets for analysis (1780 lesions), of which 11 evaluated rinsing (Allegra 2009; Cancela-Rodriguez 2011; Chaudhari 2013 (two datasets); Chen 2007; Cheng 2003b; Nagaraju 2010; Petruzzi 2014; Shukla 2018; Upadhyay 2011; Warnakulasuriya 1996), and 11 evaluated staining using a dye impregnated swab (Adil 2017; Awan 2012; Cheng 2003a; Du 2007; Jayasinghe 2020; Mashberg 1980; Onofre 2001; Rahman 2012; Silverman 1984; Singh 2015; Yamamoto 2017).
- Twenty studies evaluated oral cytology. These studies provided 24 datasets (2892 lesions) for analysis of which 16 evaluated the use of a brush to harvest cells (Delavarian 2010; Fontes 2013; Kämmerer 2013; Kaur 2016; Koch 2011a; Mehrotra 2008; Mehrotra 2011; Nanayakkara 2016; Ng 2012; Rahman 2012; Scheifele 2004; Sciubba 1999; Seijas-Naya 2012; Skandarajah 2017; Svirsky 2002; Trakroo 2015), five used DNA-image cytometry (DNA IC) (Kämmerer 2013 (two datasets); Kaur 2016; Liu 2019; Ma 2014), and three used 'scraping' (Nanayakkara 2016; Navone 2004; Navone 2008).
- Twenty-three studies evaluated light-based technologies. These studies provided 24 datasets (2587 lesions) for analysis of which 17 (Adil 2017; Amirchaghmaghi 2018; Awan 2011; Chiang 2019; Farah 2012; Ganga 2017; Hanken 2013; Koch 2011b; Leunig 2000; Mehrotra 2010; Onizawa 1999; Petruzzi 2014; Scheer 2011; Sharwani 2006a; Shi 2019; Yamamoto 2017 (two datasets)) evaluated autofluorescence, and six evaluated tissue reflectance (Awan 2012; Chainani-Wu 2015; Farah 2007; McIntosh 2009; Shukla 2018; Ujaoney 2012). One study (Sharwani 2006b) evaluated electric scattering spectroscopy.
- Nine studies evaluated index tests used in combination. These studies provided nine datasets (683 lesions) of which five assessed vital staining with light-based detection (Epstein 2008; Chainani-Wu 2015; Mehrotra 2010; Mojsa 2012; Ujaoney 2012), two assessed the use of vital staining with brush cytology (Gupta 2007; Guneri 2011), and two assessed staining method combining wheat germ agglutinin fluorescein isothiocyanate (WGA-FITC) (Baeten 2018; Johnson 2019).

All studies used a reference test of biopsy and histological examination.

The main reasons for exclusion were inappropriate patient selection or the reference standard not being applied to all patients. Further details are provided in the Characteristics of excluded studies table.

Methodological quality of included studies

Figure 2 summarises the results of the quality assessment of the included studies. Two studies (Chiang 2019; Hanken 2013) were classed as being at low risk of bias across all domains, and 33 studies were at low concern for applicability across the three domains, where patient selection, the index test, and the reference standard used were generalisable across the population attending secondary care. We recorded an unclear level of concern when a population had been selected to be made up of high or low risk patients only. Index and reference tests were classified as unclear when the description lacked sufficient detail to determine applicability of the study methods. Individual assessment for each study is provided in Figure 3.

Figure 2. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.

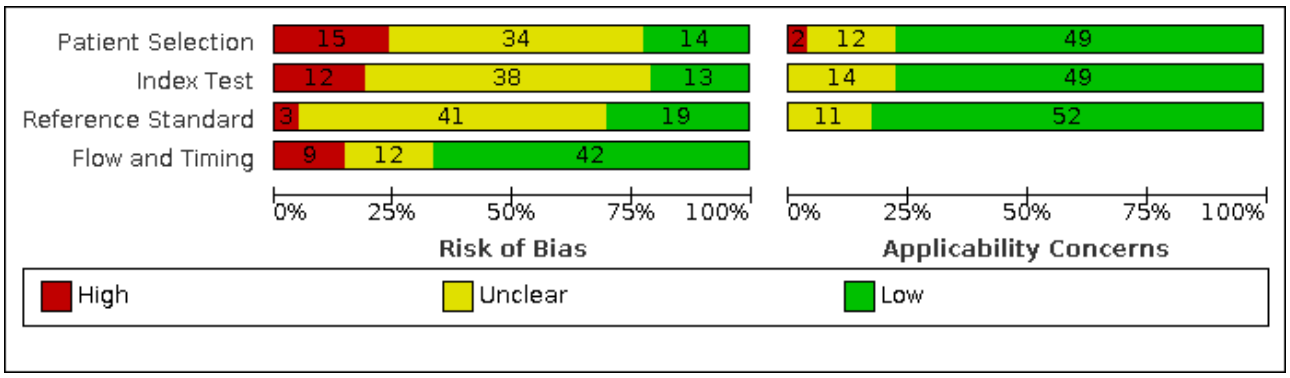


Figure 3. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

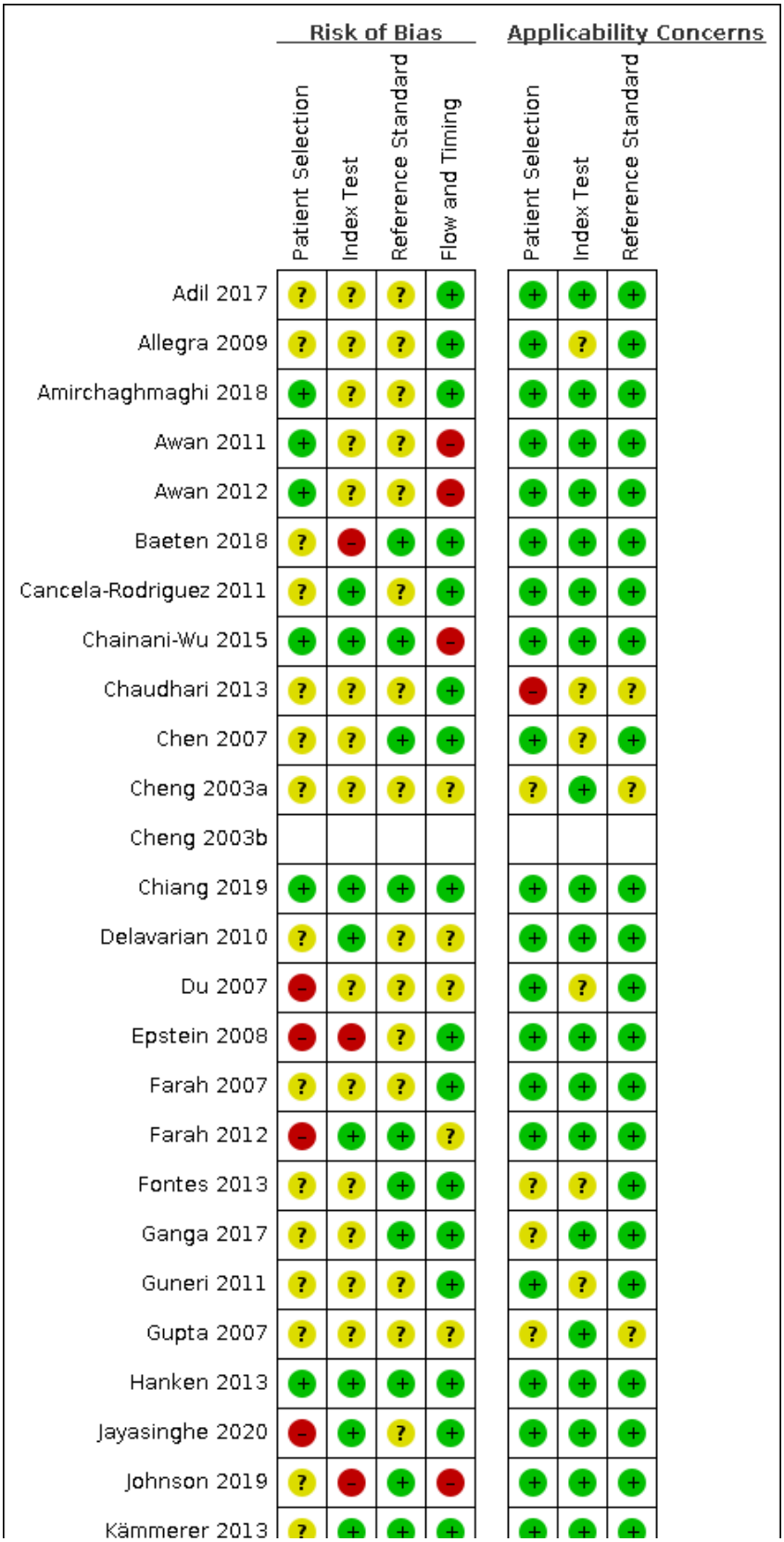


Figure 3. (Continued)

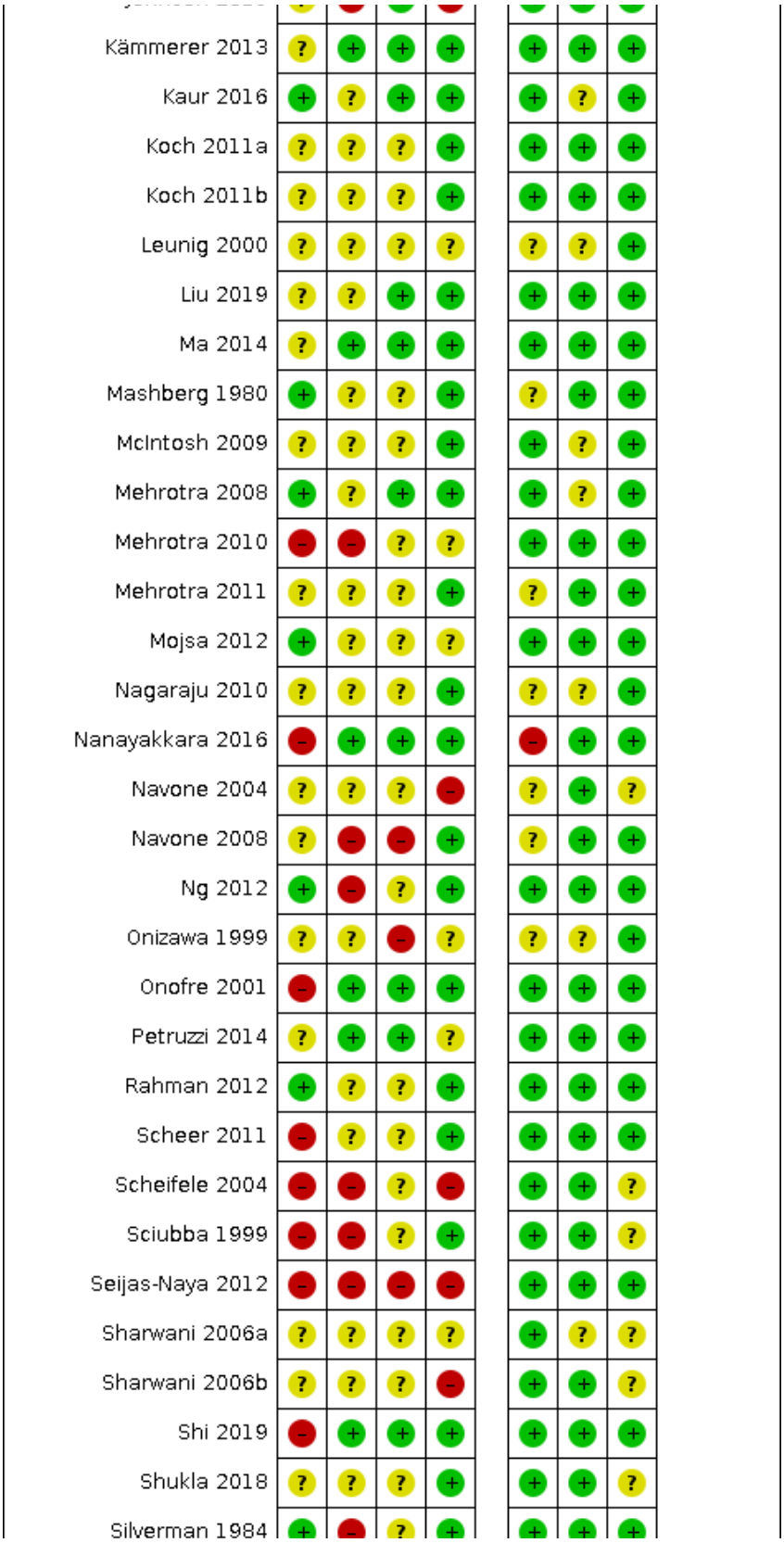
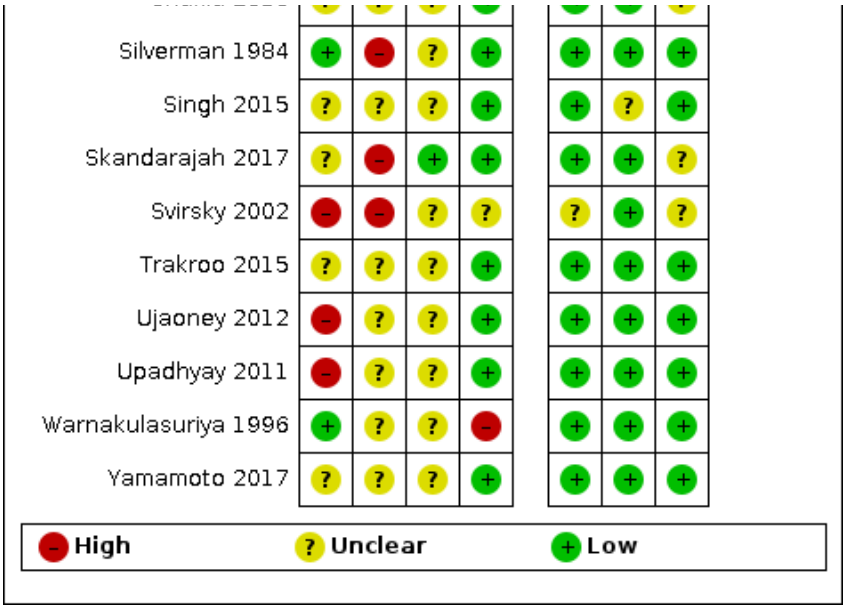


Figure 3. (Continued)



Patient selection was considered to be at low risk of bias in 14 studies, and unclear risk of bias in 34 studies. Only 16 studies reported recruiting a random or consecutive sample of patients. Recruitment was poorly reported in general, with 34 studies that failed to describe the patient selection criteria in sufficient detail and were therefore assessed as being at unclear risk of bias. 15 studies selected patients that were either at high or low risk and therefore inappropriately excluded people (Du 2007; Epstein 2008; Farah 2012; Jayasinghe 2020; Mehrotra 2010; Nanayakkara 2016; Onofre 2001; Scheer 2011; Scheifele 2004; Sciubba 1999; Seijas-Naya 2012; Shi 2019; Svirsky 2002; Ujaoney 2012; Upadhyay 2011). Most studies were judged to be at low concern for patient selection (49 studies).

The index test domain was considered to be at low risk of bias in 13 studies, and unclear risk of bias in 38 studies. The most common reason was a lack of detail about training or calibration of the clinicians. Only seven of the included studies clearly reported a training or calibration component (Cancela-Rodriguez 2011; Chaudhari 2013; Farah 2012; Hanken 2013; Nanayakkara 2016; Onofre 2001; Petruzzi 2014). In addition, a small number of studies reported the supply or development of the diagnostic aids being investigated and other study consumables (Baeten 2018; Chainani-Wu 2015; Epstein 2008; Johnson 2019; Ng 2012; Scheifele 2004; Sciubba 1999; Skandarajah 2017; Svirsky 2002; Warnakulasuriya 1996). Most studies reported the use of a pre-specified threshold (48 studies). The majority of studies were judged to be at low concern for index test (49 studies).

The reference standard domain was considered to be at low risk of bias in 19 studies and unclear in 41 studies. All of the included studies used an appropriate reference standard; a biopsy followed by a histological examination by an experienced oral pathologist. However, many did not provide sufficient details of the biopsy methods or the criteria used in the histological examination. As a consequence of lack of reported details we judged only 27 studies as being likely to correctly classify the target condition, and 31 studies as being interpreted without knowledge of the index test. The majority of studies were judged to be at low concern for the reference standard (52 studies).

The flow and timing domain was considered to be at low risk of bias in 42 studies, and unclear risk of bias in 12 studies. Most studies reported a minimal time delay between the index test and the reference standard or this could be inferred from the description of the methods. Seijas-Naya 2012 reported a three-week interval and Scheifele 2004 reported one-month interval. We classified 17 studies as unclear as they failed to explicitly report the timing between the index test and reference standard. 23 studies excluded some patients from the analysis (Awan 2011; Awan 2012; Baeten 2018; Chainani-Wu 2015; Chiang 2019; Du 2007; Epstein 2008; Guneri 2011; Jayasinghe 2020; Johnson 2019; Kaur 2016; Koch 2011a; Mehrotra 2008; Mehrotra 2011; Nanayakkara 2016; Navone 2004; Navone 2008; Ng 2012; Scheifele 2004; Sharwani 2006a; Sharwani 2006b; Shukla 2018; Warnakulasuriya 1996).

Vital staining

Patient selection was considered to be at low risk of bias in five studies (Awan 2012; Mashberg 1980; Rahman 2012; Silverman 1984; Warnakulasuriya 1996). Only four studies reported recruiting a random or consecutive sample of patients. Recruitment was poorly reported in general, with 11 studies that failed to describe the patient selection criteria in sufficient detail and were therefore assessed as being at unclear risk of bias. Four studies selected patients that were either at high or low risk and therefore inappropriately excluded people (Du 2007; Jayasinghe 2020; Onofre 2001; Upadhyay 2011). Most studies were judged to be at low concern for patient selection (16 studies).

The index test domain was considered to be at low risk of bias in 14 studies, and unclear risk of bias in six studies. Most studies reported the use of a pre-specified threshold (16 studies). The majority of studies were judged to be at low concern for index test (14 studies).

The reference standard domain was considered to be at low risk of bias in three studies (Chen 2007; Onofre 2001; Petruzzi 2014) and unclear in 17 studies. As a consequence of lack of reported details we judged only eight studies as being likely to correctly classify the target condition, and eight studies as being interpreted without knowledge of the index test. Most studies were judged to be at low concern for the reference standard (17 studies).

The flow and timing domain was considered to be at low risk of bias in 15 studies, and unclear risk of bias in three studies. Most studies reported a minimal time delay between the index test and the reference standard or this could be inferred from the description of the methods.

No studies were judged to be at low risk of bias overall, 11 studies were judged to be at low concern for applicability across all three domains.

Oral cytology

Patient selection was considered to be at low risk of bias in four studies (Kaur 2016; Mehrotra 2008; Ng 2012; Rahman 2012). Only five studies reported recruiting a random or consecutive sample of patients. Recruitment was poorly reported in general, with 11 studies that failed to describe the patient selection criteria in sufficient detail and were therefore assessed as being at unclear risk of bias. Five studies selected patients that were either at high or low risk and therefore inappropriately excluded people (Nanayakkara 2016; Scheifele 2004; Sciubba 1999; Seijas-Naya 2012; Svirsky 2002). Most studies were judged to be at low concern for patient selection (14 studies).

The index test domain was considered to be at low risk of bias in 17 studies, and at unclear risk of bias in three studies. Most studies reported the use of a pre-specified threshold (15 studies). The majority of studies were judged to be at low concern for index test (17 studies).

The reference standard domain was considered to be at low risk of bias in eight studies (Fontes 2013; Kämmerer 2013; Kaur 2016; Liu 2019; Ma 2014; Mehrotra 2008; Nanayakkara 2016; Skandarajah 2017) and unclear in 10 studies. As a consequence of lack of reported details we judged only eight studies as being likely to correctly classify the target condition, and 11 studies as being interpreted without knowledge of the index test. The majority of studies were judged to be at low concern for the reference standard (15 studies).

The flow and timing domain was considered to be at low risk of bias in 15 studies, and at unclear risk of bias in two studies. Most studies reported a minimal time delay between the index test and the reference standard or this could be inferred from the description of the methods. Seijas-Naya 2012 reported a three-week interval and Scheifele 2004 reported a one-month interval.

No studies were judged to be at low risk of bias overall, nine studies were judged to be at low concern for applicability across all three domains (Delavarian 2010; Kämmerer 2013; Koch 2011a; Liu 2019; Ma 2014; Ng 2012; Rahman 2012; Seijas-Naya 2012; Trakroo 2015).

Light-based detection

Patient selection was considered to be at low risk of bias in six studies (Amirchaghmaghi 2018; Awan 2011; Awan 2012; Chainani-Wu 2015; Chiang 2019; Hanken 2013). Only eight studies reported recruiting a random or consecutive sample of patients. Recruitment was poorly reported in general, with 12 studies that failed to describe the patient selection criteria in sufficient detail and were therefore assessed as being at unclear risk of bias. Five studies selected patients that were either at high or low risk and therefore inappropriately excluded people (Farah 2012; Mehrotra 2010; Scheer 2011; Shi 2019; Ujaoney 2012). Most studies were judged to be at low concern for patient selection (20 studies).

The index test domain was considered to be at low risk of bias in six studies (Chainani-Wu 2015; Chiang 2019; Farah 2012; Hanken 2013; Petruzzi 2014; Shi 2019), and unclear risk of bias in 16 studies. Most studies reported the use of a pre-specified threshold (19 studies). The majority of studies were judged to be at low concern for index test (19 studies).

The reference standard domain was considered to be at low risk of bias in seven studies (Chainani-Wu 2015; Chiang 2019; Farah 2012; Ganga 2017; Hanken 2013; Petruzzi 2014; Shi 2019) and unclear in 15 studies. As a consequence of lack of reported details we judged only 13 studies as being likely to correctly classify the target condition, and 11 studies as being interpreted without knowledge of the index test. Most studies were judged to be at low concern for the reference standard (20 studies).

The flow and timing domain was considered to be at low risk of bias in 13 studies, and at unclear risk of bias in six studies. Most studies reported a minimal time delay between the index test and the reference standard or this could be inferred from the description of the methods.

Two studies (Chiang 2019; Hanken 2013) were judged to be at low risk of bias overall, 16 studies were judged to be at low concern for applicability across all three domains.

Combined tests

Patient selection was considered to be at low risk of bias in two studies (Chainani-Wu 2015; Mojsa 2012). Only three studies reported recruiting a random or consecutive sample of patients. Recruitment was poorly reported in general, with four studies that failed to describe the patient selection criteria in sufficient detail and were therefore assessed as being at unclear risk of bias. Three studies selected patients that were either at high or low risk and therefore inappropriately excluded people (Epstein 2008; Mehrotra 2010; Ujaoney 2012). Most studies were judged to be at low concern for patient selection (eight studies).

The index test domain was considered to be at low risk of bias in one study (Chainani-Wu 2015), and at unclear risk of bias in four studies. Most studies reported the use of a pre-specified threshold (six studies). The majority of studies were judged to be at low concern for index test (eight studies).

The reference standard domain was considered to be at low risk of bias in three studies (Baeten 2018; Chainani-Wu 2015; Johnson 2019) and unclear in six studies with only three studies judged as being likely to correctly classify the target condition, and six studies as being interpreted without knowledge of the index test. Most studies were judged to be at low concern for the reference standard (eight studies).

The flow and timing domain was considered to be at low risk of bias in four studies, and at unclear risk of bias in three studies. Most studies reported a minimal time delay between the index test and the reference standard or this could be inferred from the description of the methods.

No studies were judged to be at low risk of bias overall, seven studies were judged to be at low concern for applicability across all three domains.

Findings

We explored the comparative accuracy of the different test types, cytology, light-based, vital staining, and combined tests (vital staining with light or cytology).

- Vital staining: 22 datasets, 1056 diseased lesions (any dysplasia or oral cavity squamous cell carcinoma (OSCC)) of a total of 1780 lesions evaluated, prevalence 59.3%.
- Oral cytology: 24 datasets, 1496 diseased lesions (any dysplasia or OSCC) of a total of 2892 lesions evaluated, prevalence 51.7%.
- Light-based: 24 datasets, 1204 diseased lesions (any dysplasia or OSCC) of a total of 2587 lesions evaluated, prevalence 46.5%.
- Vital staining plus adjunct: 9 datasets, 250 diseased lesions (any dysplasia or OSCC) of a total of 683 lesions evaluated, prevalence 36.6%.

The summary estimates for each index test were:

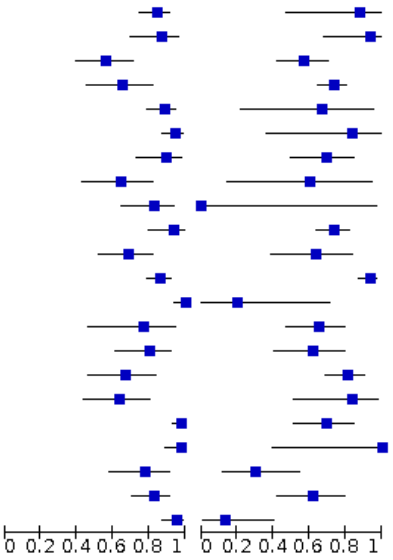
- vital staining: sensitivity 0.86 (95% confidence interval (CI) 0.79 to 0.90) specificity 0.68 (95% CI 0.58 to 0.77), 20 studies, sensitivity low-certainty evidence, specificity very low-certainty evidence;
- oral cytology: sensitivity 0.90 (95% CI 0.82 to 0.94) specificity 0.94 (95% CI 0.88 to 0.97), 20 studies, sensitivity moderate-certainty evidence, specificity moderate-certainty evidence;
- light-based: sensitivity 0.87 (95% CI 0.78 to 0.93) specificity 0.50 (95% CI 0.32 to 0.68), 23 studies, sensitivity low-certainty evidence, specificity very low-certainty evidence; and
- vital staining plus adjunct: sensitivity 0.78 (95% CI 0.45 to 0.94) specificity 0.71 (95% CI 0.53 to 0.84), 9 studies, sensitivity very low-certainty evidence, specificity very low-certainty evidence.

The estimates are displayed in the forest plots and receiver operating characteristic (ROC) plot (Figure 4; Figure 5). The confidence regions and prediction regions for cytology were relatively narrow compared to the other tests, where the prediction regions indicated a substantial amount of unobserved heterogeneity.

Figure 4. Forest plot of tests: cytology, light-based, vital staining, combined tests that included 63 studies (79 datasets) evaluating 7942 lesions (TP: true positives, FP: false posititves, FN: false negatives, TN: true negatives).

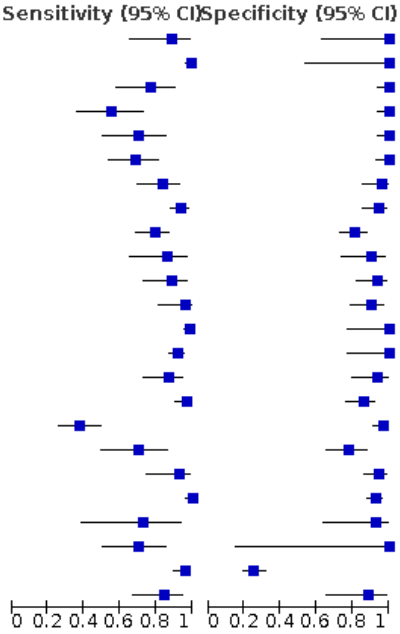
Vital staining with application method covariate

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Adil 2017	69	1	13	7	0.84 [0.74, 0.91]	0.88 [0.47, 1.00]
Allegra 2009	26	1	4	14	0.87 [0.69, 0.96]	0.93 [0.68, 1.00]
Awan 2012	23	22	18	29	0.56 [0.40, 0.72]	0.57 [0.42, 0.71]
Cancela-Rodriguez 2011	19	35	10	96	0.66 [0.46, 0.82]	0.73 [0.65, 0.81]
Chaudhari 2013	67	2	9	4	0.88 [0.79, 0.94]	0.67 [0.22, 0.96]
Chaudhari 2013	72	1	4	5	0.95 [0.87, 0.99]	0.83 [0.36, 1.00]
Chen 2007	26	9	3	20	0.90 [0.73, 0.98]	0.69 [0.49, 0.85]
Cheng 2003a	16	2	9	3	0.64 [0.43, 0.82]	0.60 [0.15, 0.95]
Cheng 2003b	24	1	5	0	0.83 [0.64, 0.94]	0.00 [0.00, 0.97]
Du 2007	31	25	2	70	0.94 [0.80, 0.99]	0.74 [0.64, 0.82]
Jayasinghe 2020	28	7	13	12	0.68 [0.52, 0.82]	0.63 [0.38, 0.84]
Mashberg 1980	101	8	16	110	0.86 [0.79, 0.92]	0.93 [0.87, 0.97]
Nagaraju 2010	55	4	0	1	1.00 [0.94, 1.00]	0.20 [0.01, 0.72]
Onofre 2001	10	13	3	24	0.77 [0.46, 0.95]	0.65 [0.47, 0.80]
Petruzzi 2014	24	10	6	16	0.80 [0.61, 0.92]	0.62 [0.41, 0.80]
Rahman 2012	18	11	9	48	0.67 [0.46, 0.83]	0.81 [0.69, 0.90]
Shukla 2018	19	2	11	10	0.63 [0.44, 0.80]	0.83 [0.52, 0.98]
Silverman 1984	97	10	2	23	0.98 [0.93, 1.00]	0.70 [0.51, 0.84]
Singh 2015	45	0	1	4	0.98 [0.88, 1.00]	1.00 [0.40, 1.00]
Upadhyay 2011	21	14	6	6	0.78 [0.58, 0.91]	0.30 [0.12, 0.54]
Warnakulasuriya 1996	47	11	10	18	0.82 [0.70, 0.91]	0.62 [0.42, 0.79]
Yamamoto 2017	61	13	3	2	0.95 [0.87, 0.99]	0.13 [0.02, 0.40]



Cytology with covariate

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Delavarian 2010	16	0	2	8	0.89 [0.65, 0.99]	1.00 [0.63, 1.00]
Fontes 2013	157	0	1	6	0.99 [0.97, 1.00]	1.00 [0.54, 1.00]
Kämmerer 2013	23	0	7	57	0.77 [0.58, 0.90]	1.00 [0.94, 1.00]
Kämmerer 2013	17	0	14	57	0.55 [0.36, 0.73]	1.00 [0.94, 1.00]
Kämmerer 2013	21	0	9	57	0.70 [0.51, 0.85]	1.00 [0.94, 1.00]
Kaur 2016	33	0	15	48	0.69 [0.54, 0.81]	1.00 [0.93, 1.00]
Kaur 2016	40	2	8	46	0.83 [0.70, 0.93]	0.96 [0.86, 0.99]
Koch 2011a	108	4	7	63	0.94 [0.88, 0.98]	0.94 [0.85, 0.98]
Liu 2019	68	22	18	95	0.79 [0.69, 0.87]	0.81 [0.73, 0.88]
Ma 2014	19	3	3	27	0.86 [0.65, 0.97]	0.90 [0.73, 0.98]
Mehrotra 2008	30	3	4	42	0.88 [0.73, 0.97]	0.93 [0.82, 0.99]
Mehrotra 2011	26	5	1	47	0.96 [0.81, 1.00]	0.90 [0.79, 0.97]
Nanayakkara 2016	165	0	2	14	0.99 [0.96, 1.00]	1.00 [0.77, 1.00]
Nanayakkara 2016	154	0	13	14	0.92 [0.87, 0.96]	1.00 [0.77, 1.00]
Navone 2004	39	2	6	31	0.87 [0.73, 0.95]	0.94 [0.80, 0.99]
Navone 2008	71	12	2	73	0.97 [0.90, 1.00]	0.86 [0.77, 0.92]
Ng 2012	27	3	45	96	0.38 [0.26, 0.50]	0.97 [0.91, 0.99]
Rahman 2012	19	13	8	46	0.70 [0.50, 0.86]	0.78 [0.65, 0.88]
Scheifele 2004	24	4	2	66	0.92 [0.75, 0.99]	0.94 [0.86, 0.98]
Sciubba 1999	102	14	0	182	1.00 [0.96, 1.00]	0.93 [0.88, 0.96]
Seijas-Naya 2012	8	1	3	12	0.73 [0.39, 0.94]	0.92 [0.64, 1.00]
Skandarajah 2017	21	0	9	2	0.70 [0.51, 0.85]	1.00 [0.16, 1.00]
Svirsky 2002	93	150	4	51	0.96 [0.90, 0.99]	0.25 [0.20, 0.32]
Trakroo 2015	27	2	5	16	0.84 [0.67, 0.95]	0.89 [0.65, 0.99]



Light-based with covariate

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Adil 2017	70	2	12	6	0.85 [0.76, 0.92]	0.75 [0.35, 0.97]
Amirchaghmaghi 2018	21	29	0	4	1.00 [0.84, 1.00]	0.12 [0.03, 0.28]
Awan 2011	37	61	7	11	0.84 [0.70, 0.93]	0.15 [0.08, 0.26]
Awan 2012	34	52	10	20	0.77 [0.62, 0.89]	0.28 [0.18, 0.40]
Chainani-Wu 2015	2	13	15	40	0.12 [0.01, 0.36]	0.75 [0.62, 0.86]
Chiang 2019	56	32	7	25	0.89 [0.78, 0.95]	0.44 [0.31, 0.58]
Farah 2007	10	45	0	0	1.00 [0.69, 1.00]	0.00 [0.00, 0.08]
Farah 2012	11	29	15	63	0.42 [0.23, 0.63]	0.68 [0.58, 0.78]
Ganga 2017	19	59	6	116	0.76 [0.55, 0.91]	0.66 [0.59, 0.73]
Hanken 2013	47	8	1	4	0.98 [0.89, 1.00]	0.33 [0.10, 0.65]
Koch 2011b	31	1	2	44	0.94 [0.80, 0.99]	0.98 [0.88, 1.00]

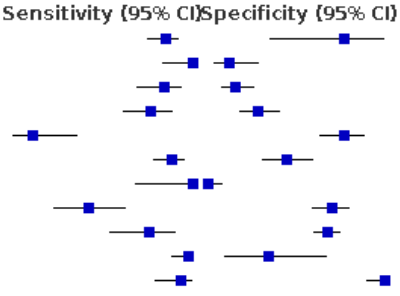
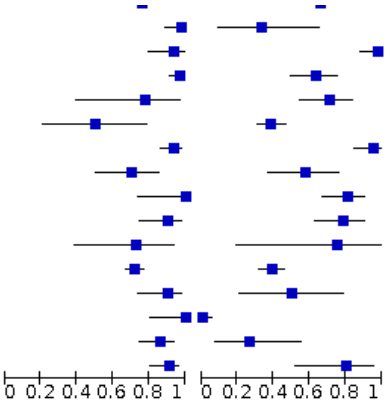


Figure 4. (Continued)

Hanken 2013	47	8	1	4	0.98 [0.89, 1.00]	0.33 [0.10, 0.65]
Koch 2011b	31	1	2	44	0.94 [0.80, 0.99]	0.98 [0.88, 1.00]
Leunig 2000	97	22	3	38	0.97 [0.91, 0.99]	0.63 [0.50, 0.75]
McIntosh 2009	7	12	2	29	0.78 [0.40, 0.97]	0.71 [0.54, 0.84]
Mehrotra 2010	6	88	6	56	0.50 [0.21, 0.79]	0.39 [0.31, 0.47]
Onizawa 1999	75	2	5	42	0.94 [0.86, 0.98]	0.95 [0.85, 0.99]
Petruzzi 2014	21	11	9	15	0.70 [0.51, 0.85]	0.58 [0.37, 0.77]
Scheer 2011	12	10	0	42	1.00 [0.74, 1.00]	0.81 [0.67, 0.90]
Sharwani 2006a	28	8	3	30	0.90 [0.74, 0.98]	0.79 [0.63, 0.90]
Sharwani 2006b	8	1	3	3	0.73 [0.39, 0.94]	0.75 [0.19, 0.99]
Shi 2019	239	113	92	73	0.72 [0.67, 0.77]	0.39 [0.32, 0.47]
Shukla 2018	27	6	3	6	0.90 [0.73, 0.98]	0.50 [0.21, 0.79]
Ujaoney 2012	17	81	0	1	1.00 [0.80, 1.00]	0.01 [0.00, 0.07]
Yamamoto 2017	55	11	9	4	0.86 [0.75, 0.93]	0.27 [0.08, 0.55]
Yamamoto 2017	58	3	6	12	0.91 [0.81, 0.96]	0.80 [0.52, 0.96]



Vital staining plus adjunct (light or cytology)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Baeten 2018	32	4	4	13	0.89 [0.74, 0.97]	0.76 [0.50, 0.93]
Chainani-Wu 2015	2	7	15	46	0.12 [0.01, 0.36]	0.87 [0.75, 0.95]
Epstein 2008	30	9	12	28	0.71 [0.55, 0.84]	0.76 [0.59, 0.88]
Guneri 2011	14	22	1	6	0.93 [0.68, 1.00]	0.21 [0.08, 0.41]
Gupta 2007	44	3	2	47	0.96 [0.85, 0.99]	0.94 [0.83, 0.99]
Johnson 2019	55	9	11	25	0.83 [0.72, 0.91]	0.74 [0.56, 0.87]
Mehrotra 2010	0	24	4	74	0.00 [0.00, 0.60]	0.76 [0.66, 0.84]
Mojsa 2012	7	25	0	9	1.00 [0.59, 1.00]	0.26 [0.13, 0.44]
Ujaoney 2012	10	17	7	65	0.59 [0.33, 0.82]	0.79 [0.69, 0.87]

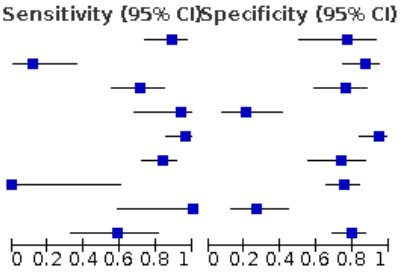
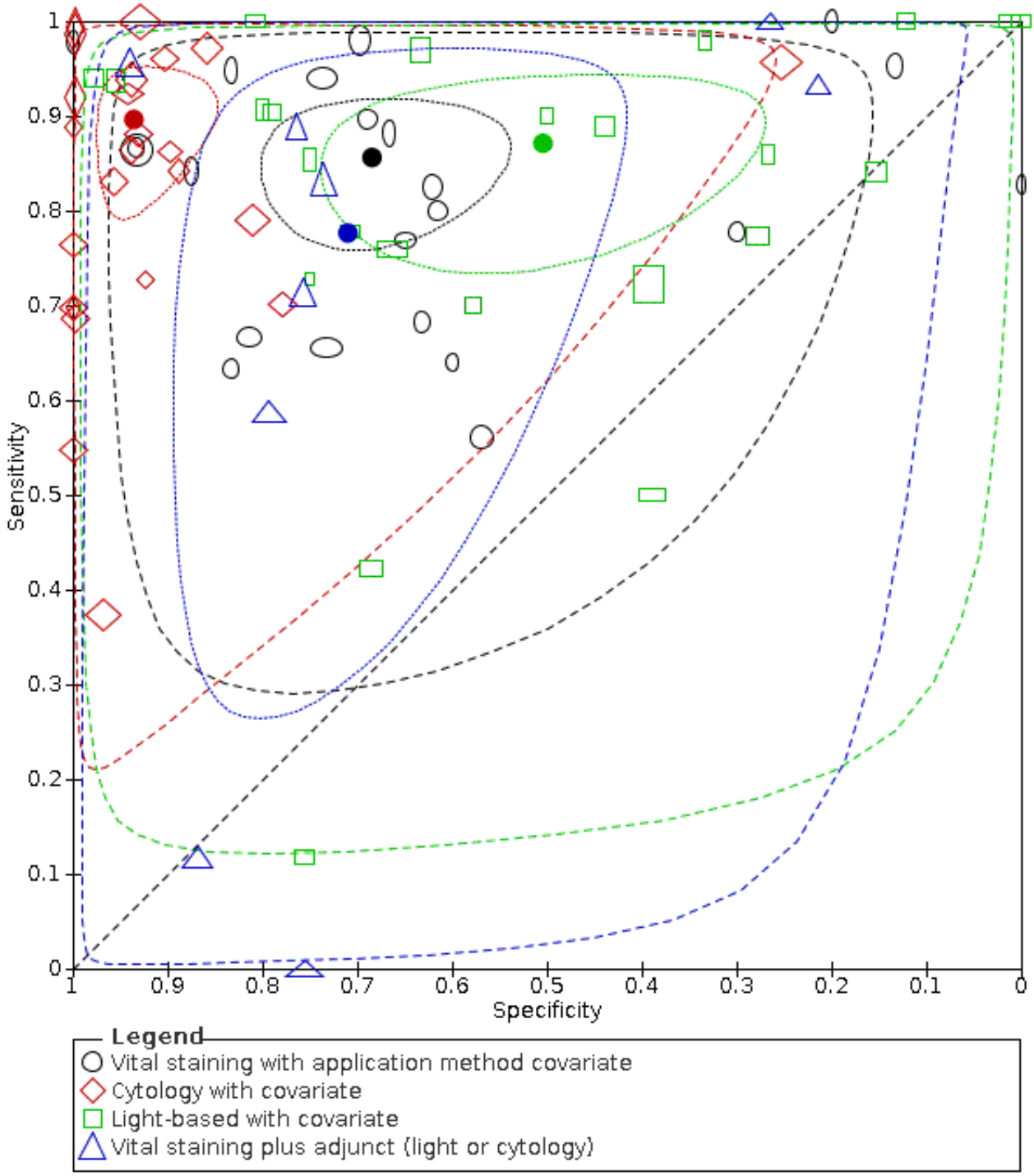


Figure 5. Summary receiver operating characteristic (SROC) plot of tests: cytology, light-based, vital staining, and combined tests.



We undertook statistical analysis of the differences in sensitivity and specificity across the index tests to determine the sources of differences. Using vital staining as the reference category we observed the following:

- difference in sensitivity for oral cytology -0.04 (-0.12 to 0.04) P = 0.33, light-based -0.02 (-0.11 to 0.08) P = 0.70, combined tests 0.08 (-0.18 to 0.33);
- difference in specificity for oral cytology -0.25 (-0.36 to -0.14) P < 0.001, light-based 0.18 (-0.03 to 0.39) P = 0.10, combined tests -0.03 (-0.21 to 0.16)

Results of the relative accuracy of the tests are presented in Summary of findings 1.

Ideally, we would have performed direct test comparisons based on studies that compared more than one index test on the same sample. However, only seven studies directly compared index tests in the different test categories on the same sample: light-based versus staining (Adil 2017; Petruzzi 2014; Shukla 2018; Yamamoto 2017), staining versus cytology (Rahman 2012), or single tests versus combined tests (Chainani-Wu 2015; Ujaoney 2012). Five other studies compared index tests within the same categories (Awan 2011; Chaudhari 2013; Kämmerer 2013; Kaur 2016; Nanayakkara 2016).

For vital staining, we graded the certainty of the evidence for sensitivity as low, downgrading one level for study limitations (risk of bias for the patient selection domain) and one level for inconsistency; for specificity we graded the evidence as very low, additionally downgrading for imprecision. For oral cytology we graded the evidence as moderate for sensitivity and specificity, downgrading once for study limitations (risk of bias for the index test domain). For light-based index tests, we graded the certainty of the evidence for sensitivity as low, downgrading one level for study limitations (risk of bias for the patient selection domain) and one level for inconsistency; for specificity we graded the evidence as very low, additionally downgrading for imprecision. For the combined index tests, we graded the certainty of the evidence for sensitivity and specificity as very low, downgrading one level for study limitations (risk of bias for the patient selection and index test domains), imprecision, and inconsistency.

Investigations of heterogeneity

Vital staining

Four datasets had a prevalence of oral potentially malignant disorders (OPMD) or OSCC < 35%, 18 datasets had a prevalence greater or equal to 35%. We explored the effect of disease prevalence (< 35%) in the primary studies on sensitivity or specificity or both for the bivariate model through meta-regression. The log likelihood for the base model assuming equality of variances but without covariates was -126.53. Adding covariates for test type on sensitivity and specificity resulted in a log likelihood for this most complex model of -124.94 (likelihood ratio (LR) test Chi^2 1.19, degrees of freedom (df) 2, P = 0.55). Relaxing the assumption of equal variances resulted in the following estimates:

- low prevalence in the study sample (< 35%): sensitivity 0.78 (95% CI 0.61 to 0.89) specificity 0.74 (95% CI 0.69 to 0.78);
- higher prevalence in the study sample (>= 35%): sensitivity 0.87 (95% CI 0.79 to 0.92) specificity 0.67 (95% CI 0.52 to 0.79)

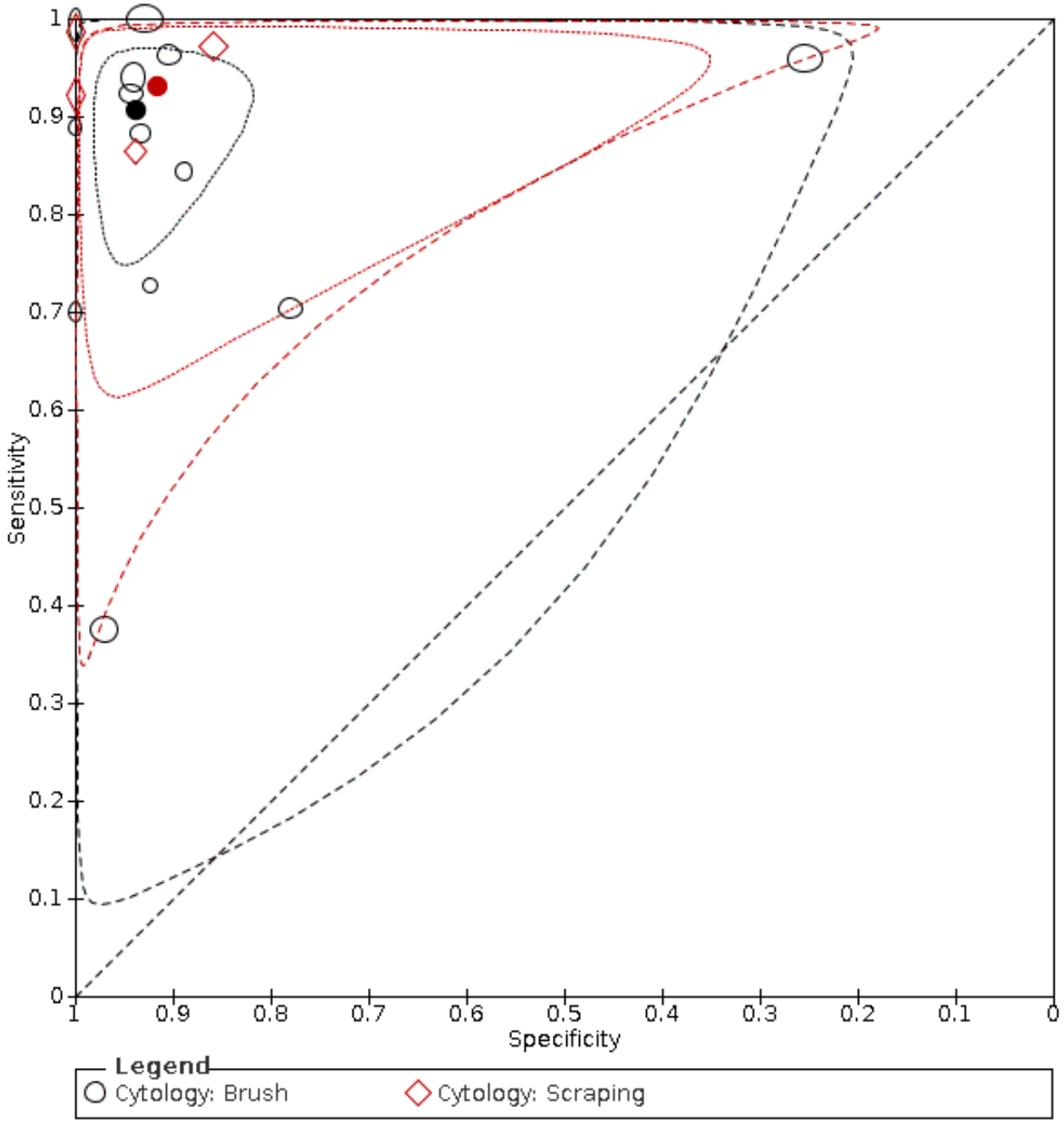
with insufficient evidence of a difference in sensitivity or specificity (-0.09 (95% CI -0.24 to 0.06) and 0.07 (95% CI -0.07 to 0.22)).

Oral cytology

Twenty studies provided 24 datasets for analysis of which 16 evaluated the use of a brush to harvest cells (Delavarian 2010; Fontes 2013; Kämmerer 2013; Kaur 2016; Koch 2011a; Mehrotra 2008; Mehrotra 2011; Nanayakkara 2016; Ng 2012; Rahman 2012; Scheifele 2004; Sciubba 1999; Seijas-Naya 2012; Skandarajah 2017; Svirsky 2002; Trakroo 2015) with 995 diseased lesions from a total of 1950 lesions evaluated, five used DNA-image cytometry (DNA-IC) (Kämmerer 2013; Kaur 2016 (two datasets); Liu 2019; Ma 2014) with 216 diseased lesions evaluated from a total of 525, and three used 'scraping' (Nanayakkara 2016; Navone 2004; Navone 2008) with 285 lesions evaluated from a total of 417. We explored potential differences in diagnostic test accuracy through meta-regression according to the different cytology methods used.

The log likelihood for the base model without covariates was -132.89. Adding covariates according to the different methodology resulted in a log likelihood for this most complex model of -131.61, a negligible improvement to the model (LR test Chi^2 5.72, df 4, P = 0.22). Relaxing the assumption of equal variances (LR test Chi^2 14.79, df 6, P = 0.02) resulted in the estimates given below and illustrated in Figure 6:

Figure 6. Summary receiver operating characteristic (SROC) plot of cytology grouped by different techniques.



- brush: sensitivity 0.91 (95% CI 0.81 to 0.96) specificity 0.94 (95% CI 0.87 to 0.97);
- DNA-IC: sensitivity 0.76 (95% CI 0.68 to 0.82) specificity 0.98 (95% CI 0.72 to 0.99);
- scraping: sensitivity 0.93 (95% CI 0.87 to 0.96) specificity 0.92 (95% CI 0.81 to 0.97).

The 95% prediction regions indicate variability across the tests, with the greatest unexplained heterogeneity for the use of a brush to harvest the cells.

Seven datasets had a prevalence of OPMD or OSCC < 35%, 17 datasets had a prevalence greater or equal to 35%. We explored the effect of disease prevalence (< 35%) in the primary studies on sensitivity or specificity or both for the bivariate model through meta-regression. The log likelihood for the base model assuming equality of variances but without covariates was -132.89. Adding covariates for test type on sensitivity and specificity resulted in a log likelihood for this most complex model of -130.04 (LR test Chi^2 5.69, df 2, P = 0.06). Relaxing the assumption of equal variances resulted in the following estimates:

- low prevalence in the study sample (< 35%): sensitivity 0.94 (95% CI 0.80 to 0.98) specificity 0.91 (95% CI 0.63 to 0.98);
- higher prevalence in the study sample (>= 35%): sensitivity 0.87 (95% CI 0.77 to 0.93) specificity 0.95 (95% CI 0.90 to 0.97)

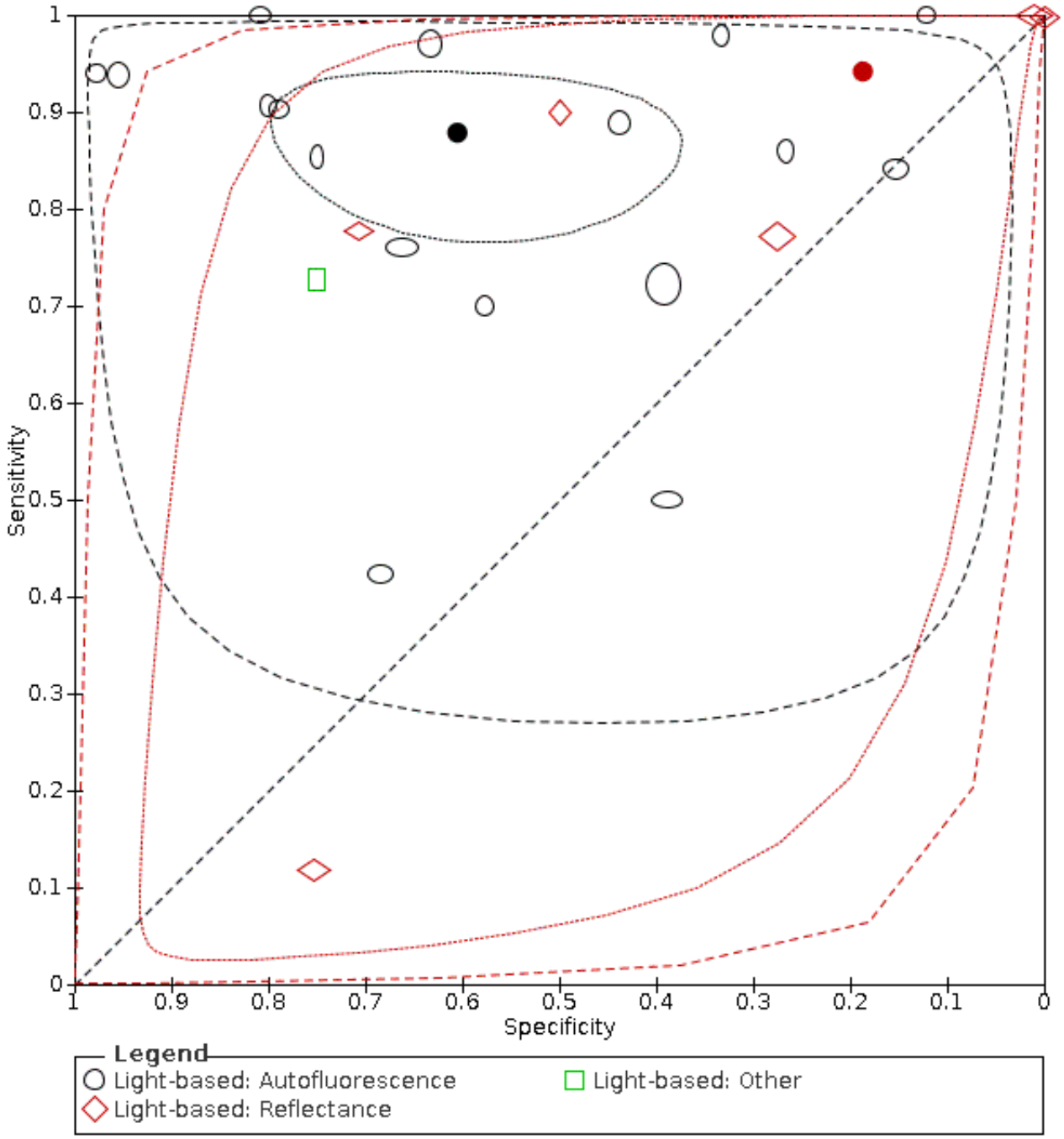
with insufficient evidence of a difference in sensitivity or specificity (0.07 (95% CI -0.04 to 0.18) and -0.04 (95% CI -0.20 to 0.11).

Light-based detection

Twenty-three studies provided 24 datasets for analysis of which 17 (Adil 2017; Amirchaghmaghi 2018; Awan 2011; Chiang 2019; Farah 2012; Ganga 2017; Hanken 2013; Koch 2011b; Leunig 2000; Mehrotra 2010; Onizawa 1999; Petruzzi 2014; Scheer 2011; Sharwani 2006a; Shi 2019; Yamamoto 2017 (two datasets)) evaluated autofluorescence, with 1066 diseased lesions from a total of 2140 lesions evaluated, and six evaluated tissue reflectance (Awan 2012; Chainani-Wu 2015; Farah 2007; McIntosh 2009; Shukla 2018; Ujaoney 2012) with 127 diseased lesions from a total of 432 lesions evaluated. One study (Sharwani 2006b) evaluated electric scattering spectroscopy with 11 diseased lesions from a total of 15 lesions evaluated. We explored potential differences in diagnostic test accuracy through meta-regression with method of fluorescence, autofluorescence or tissue reflectance, as a covariate. For the purpose of this analysis we excluded the single study which evaluated electric scattering spectroscopy (Sharwani 2006b).

The log likelihood for the base model assuming equality of variances but without covariates was -154.93. Adding covariates for method of fluorescence resulted in a log likelihood for this most complex model of -152.54, a negligible improvement to the model (LR test Chi^2 4.78, df 2, P = 0.08). Relaxing the assumption of equal variances (LR test Chi^2 10.99, df 3, P = 5) resulted in the estimates given below and illustrated in Figure 7:

Figure 7. Summary receiver operating characteristic (SROC) plot of light-based index tests grouped by different techniques.



- autofluorescence: sensitivity 0.88 (95% CI 0.80 to 0.93) specificity 0.61 (95% CI 0.44 to 0.75);
- tissue reflectance: sensitivity 0.94 (95% CI 0.35 to 0.99) specificity 0.19 (95% CI 0.03 to 0.67).

Eight datasets had a prevalence of OPMD or OSCC < 35%, 15 datasets had a prevalence greater or equal to 35%. We explored the effect of disease prevalence (< 35%) in the primary studies on sensitivity or specificity or both for the bivariate model through meta-regression. The log likelihood for the base model assuming equality of variances but without covariates was -154.93. Adding covariates for test type on sensitivity and specificity resulted in a log likelihood for this most complex model of -152.81 (LR test Chi^2 4.23, df 2, P = 0.12). Relaxing the assumption of equal variances resulted in the following estimates:

- low prevalence in the study sample (< 35%): sensitivity 0.88 (95% CI 0.36 to 0.99) specificity 0.36 (95% CI 0.09 to 0.76);
- higher prevalence in the study sample (>= 35%): sensitivity 0.89 (95% CI 0.84 to 0.93) specificity 0.56 (95% CI 0.37 to 0.74)

with insufficient evidence of a difference in sensitivity or specificity (-0.02 (95% CI -0.30 to 0.26) and -0.20 (95% CI -0.65 to 0.24)).

Combined tests

Four datasets had a prevalence of OPMD or OSCC < 35% and five datasets had a prevalence greater or equal to 35%. We planned to explored the effect of disease prevalence (< 35%) in the primary studies on sensitivity or specificity or both for the bivariate model through meta-regression but due to the small number of studies we were unable to do so.

DISCUSSION

Summary of main results

Our review aimed to estimate the accuracy of index tests as an adjunct to the conventional oral examination, for the detection of oral cancer or oral potentially malignant disorders (OPMDs) in patients presenting with clinically evident suspicious or innocuous lesions.

The main findings of the review are as follows.

- **Accuracy estimates differed according to test type:** oral cytology: sensitivity 0.90 (95% confidence interval (CI) 0.82 to 0.94) specificity 0.94 (95% CI 0.88 to 0.97); light-based: sensitivity 0.87 (95% CI 0.78 to 0.93) specificity 0.50 (95% CI 0.32 to 0.68); vital staining: sensitivity 0.86 (95% CI 0.79 to 0.90) specificity 0.68 (95% CI 0.58 to 0.77); vital staining plus adjunct: sensitivity 0.78 (95% CI 0.45 to 0.94) specificity 0.71 (95% CI 0.53 to 0.84). The summary estimates for the sensitivity of the tests were largely similar but the specificities differed markedly (Summary of findings 1). Furthermore, the specificity estimate of tissue reflectance was far lower than autofluorescence (autofluorescence specificity 0.61 (95% CI 0.44 to 0.75); tissue reflectance specificity 0.19 (95% CI 0.03 to 0.67)).
- For oral cytology we graded the evidence as moderate for sensitivity and specificity, downgrading once for study limitations. For light-based tests we graded the certainty of the evidence for sensitivity as low, downgrading one level for study limitations and one level for inconsistency; for specificity we graded the evidence as very low, additionally downgrading for imprecision. For vital staining we graded the certainty of the evidence for sensitivity as low, downgrading one level for study limitations and one level for inconsistency; for specificity we graded the evidence as very low, additionally downgrading for imprecision. For the combined tests we graded the certainty of the evidence for sensitivity and specificity as very low, downgrading one level for study limitations, imprecision, and inconsistency.

Strengths and weaknesses of the review

This Cochrane Review builds upon the systematic reviews conducted by the American Dental Association (ADA) (Lingen 2017) which limited its evaluation to the consideration of tests marketed to the US general dentist and by Tiwari et al who evaluated the use of optical fluorescence imaging in clinical evaluation, risk assessment or management of OPMD or oral cavity squamous cell carcinoma (OSCC) or both (Tiwari 2019). The opportunity to expand on tests performed globally and to explore novel tests, along with establishing the comparative test accuracy and investigation of potential sources of heterogeneity are key strengths of this review. A further strength is the large number of included studies (23 additional studies published since the last version of the review) and the large amount of data that were evaluated. Included studies evaluated three diagnostic tests used alone and in combination. We were therefore able to evaluate the accuracy of the different categories of index test to estimate summary values of sensitivity and specificity with an acceptable degree of precision, and to determine the relative diagnostic test accuracy between the different index tests. The included studies were diverse in geographic location, prevalence of disease (from 4% to 97%), and relative skill level of practitioners, and were therefore representative of a wide variety of clinical contexts. However, most studies were undertaken in secondary care facilities.

There were limitations in the study design and conduct of many studies, with only two assessed as being at low risk of bias across all domains. We had considered a post hoc sensitivity analysis restricting the analyses to studies assessed as at low risk of bias overall to determine the effect of risk of bias on accuracy estimates, but due to the very low number of studies satisfying this criterion we were unable to do so. Within the different categories of index test we observed considerable heterogeneity according to the prevalence of the target condition and study eligibility criteria, and also unexplained heterogeneity of study results.

The positivity threshold that was most consistently reported and therefore applied in this review is where OSCC and any form of dysplasia are treated as a positive diagnosis. The review group would urge further research to explore whether this diagnostic threshold is commensurate with our evolving understanding of carcinogenesis as a non-linear process and that grade of dysplasia may not be reliably predictive of malignant transformation. As the target condition can affect a broad area of oral mucosa, differences between the index test and reference standard could also have been caused by differences in the location where the respective tests were performed. In many of the vital staining and light-based studies, it was not clear how clinicians determined the most appropriate location of the biopsy. In some of the studies, the basal and transepithelial cell harvesting and potential ulceration caused during procurement of a cytological sample might also have made it challenging to undertake the reference standard (surgical or punch biopsy) in the same location as the index test.

Applicability of findings to the review question

There were few concerns regarding the applicability of individual studies to address the review question, however the review was dominated by studies conducted in secondary care. Whilst this is appropriate for the review question, the limited scope of the included studies in terms of clinical setting limits the applicability of the findings to this context. It is entirely reasonable to assume that the diagnostic accuracy would be different if these technologies were performed in a primary care setting by non-experts (i.e. more false positives would be generated leading to unnecessary referrals, more false negatives could also be the case if tests were not performed correctly, and it is also possible that the tests would be robust enough to be the same in general and specialist practice). Therefore it would be unwise to make inferences regarding the

performance of these devices by a frontline clinician. For example, at a general dental practitioner's office, a general dentist may reasonably be expected to see 24 OPMDs per month (Ogden 2015), of which almost all will be non-dysplastic and those that are will likely show mild dysplasia. The variation in prevalence levels impact substantially on the number of false-positive and false-negative lesions.

The findings are generalisable to the wider population due to the geographical diversity of the included studies with one proviso being that no studies originated in Africa. Furthermore, the studies focus on the performance of the adjunctive tests in a secondary care facility, so the findings should be treated with caution when considering primary care.

AUTHORS' CONCLUSIONS

Implications for practice

The lower specificity of the light-based and vital staining index tests compared to the high specificity of cytology could be explained by the underlying mechanisms rendering index positive outcomes. Light-based adjuncts, such as tissue autofluorescence devices, are designed to detect deviations from the normal absorption/emission spectra of natural fluorophores within the oral mucosa (i.e. epithelium and structures within underlying connective tissue). Loss of native fluorescence (fluorescence visualization loss (FVL)) when exposed to light in the blue spectrum is not specific to the abnormal molecular/architectural perturbations seen in dysplasia and oral cavity squamous cell carcinoma (OSCC). Other benign disease processes, particularly inflammatory mucosal diseases which can clinically mimic oral potentially malignant disorders (OPMDs), will lead to false-positive light-based test outcomes. Similarly, the underlying mechanisms of vital stains, which are largely speculative, are also not specific for dysplasia and OSCC. As an example, toluidine blue will bind to ulceration. Outcome measures for light-based and vital staining tests have not been standardised across these studies. There is too much 'noise' associated with light-based and vital staining adjuncts that will confuse the non-expert clinician. Filtering the 'noise' by gaining an appreciation for differentiating benign mimicker 'confounder' lesions from OPMDs could bolster the specificity, but studies assessing these technologies in the hands of frontline clinicians are lacking. The utility of these adjuncts should be confined for use by expert clinicians to facilitate biopsy site selection in heterogeneous OPMDs, or as a surveillance tool. In contrast, the high specificity of cytology we report should not be a surprise as cytology is surrogate for the reference standard of histopathology. Yet, current cytologic tests lack the architectural context of histopathology and outcomes cannot discriminate between grades of epithelial dysplasia and OSCC, therefore still necessitating a tissue biopsy to reach a definitive diagnosis. As such, experts would usually not employ cytology as a replacement for tissue biopsy, at least not for a baseline assessment. The value of cytology as a surveillance test by expert clinicians in high-risk patients with a history of epithelial dysplasia or OSCC for whom multiple serial biopsies are problematic has not been adequately studied. The question remains, could cytology have utility in the hands of non-expert frontline clinicians? Unfortunately, due to the paucity of studies in a primary care setting, our review cannot adequately answer this question. The prevalence of OPMDs in the general population is relatively high compared to the known incidence of OSCC (Petti 2003) and therefore most of the OPMDs detected likely represent low-grade disease with a low risk for malignant transformation. Frontline clinicians are therefore more likely to encounter low-risk lesions rather than lesions that represent high-grade dysplasias or OSCC (Ogden 2015; Purkayastha 2018).

Given the relatively high sensitivity of cytology, the triage of low-risk lesions in the general population with a non-invasive cytological test indicated for lesions that are small enough to be adequately sampled seems reasonable, although with some provisos. Frontline clinicians must appreciate that cytologic tests are not indicated for frankly suspicious malignancies or for OPMDs which display ominous high-risk clinical features (e.g. induration, pain, ulceration or large lesions (> 2 cm in diameter) or both, with heterogeneous topography (i.e. red or mixed red and white or verrucous components) suggesting the possibility of variable histopathology within the lesion field and increased risk for sampling errors by a single brush. Such lesions are better referred to experts. Furthermore, they must be cognisant that cytology is imperfect and that both 'false positives' and 'false negatives' are possible, therefore patient follow-up and repeat sampling may be necessary. Studies investigating the cost-effectiveness of cytology compared to referral have not been performed. This is particularly relevant in resource-poor countries where the role of non-invasive cytologic tests could be performed by non-clinicians.

At this point in time, none of the adjunctive tests can be recommended as a replacement for the current standard of a surgical or scalpel biopsy and histological assessment. Yet, the performance of cytology compared to histopathology shows promise; however, there is only limited evidence of the value of vital staining and cytology combined. Patients should be referred to clinicians with appropriate interest and training in the management of OPMDs, including oral medicine, oral and maxillofacial pathology, oral and maxillofacial surgery, stomatology, and head and neck surgery.

The index tests were conducted in secondary care with trained and experienced specialists; no current evidence exists for their use in primary care (Epstein 2008; Kerr 2020). Vital staining and light-based tests are dependent on the visual assessment of the lesion (Mehrotra 2010) and cytology requires adequate training and experience of performing a transepithelial biopsy for harvesting basal cells, though important relevant diagnostic information may be deduced from suprabasal cells (Ogden 1994).

Implications for research

Diagnostic test accuracy (DTA) studies addressing this area of research would be welcomed, particularly those that report on a patient basis and that report their findings according to the STARD checklist (Bossuyt 2003). As with studies of comparative effectiveness, all new DTA studies should register and publish the study protocol, and ensure that they fully address the domains within the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool (Whiting 2011).

Given the relatively high values of the summary estimates of sensitivity and specificity, cytology would appear to offer the most potential. Combined adjunctive tests involving cytology would warrant further investigation, only two studies investigated vital staining plus cytology (Guneri 2011; Gupta 2007), and no studies currently combine cytology with a light-based test. Studies should endeavour to publish test results in as much detail as possible, the

following categories are suggested as a minimum for the histology reference standard: benign, mild, moderate, severe dysplasia, and OSCC. This would allow further analysis to be undertaken to assess the severity of any misdiagnosis; if an OSCC is misclassified as a mild dysplasia by the index test then the consequences are significantly more severe than if a moderate dysplasia be classified as mild.

One area of analysis worthy of further investigation is to consider the relative performance of these technologies as the disease threshold for the target condition changes. The numbers of true positives, true negatives, false positives, and false negatives should be clearly reported for different grades of dysplasia in future studies. The threshold for this review is benign/any dysplasia or OSCC, hence the relatively high prevalence values reported in many of the included studies. Where reported, the distribution of dysplasia grades was skewed towards mild dysplasia, with typically many more cases of mild dysplasia than moderate or severe dysplasia. Because of the significantly higher risk for malignant progression of higher grade dysplasia (moderate or severe) versus lower grade dysplasia (mild), it would be informative to explore test performance at different thresholds. Knowledge of the comparative test performance at a positivity threshold of mild dysplasia (i.e. benign/mild or moderate or severe dysplasia or OSCC) versus a positivity threshold of moderate dysplasia (mild/moderate or severe dysplasia or OSCC) would be extremely useful from a clinical decision-making perspective, in terms of identifying 'the optimum' thresholds for these tests to minimise either false positives or false negatives with respect to the target conditions of interest. As an example, Ng 2012 evaluated the test accuracy of quantitative cytology with 171 lesions of which 99 (58%) were confirmed as benign. Of the dysplastic lesions 44 (28%) were considered as low risk (mild or moderate epithelial dysplasia and verrucous hyperplasia), 12 (7%) were reported as high risk (severe epithelial dysplasia and carcinoma in situ), and 16 (9%) were reported as OSCC. When the positivity threshold was set at low risk and above in accordance with our target condition, the sensitivity of Oral Advance was 0.38, the lowest of all the included studies, but applying a positivity threshold of high risk and above resulted in a sensitivity of 0.89 with specificity of 0.97, a highly acceptable level of performance for the detection of serious lesions. In contrast, raising the threshold to capture serious disease comes with the potential to miss OPMDs with mild or moderate dysplasia harbouring genetic/epigenetic alterations predicting a high risk for progression to malignancy. Threshold considerations must be applied to each technology. Meta-analysis methods that deal with multiple categories within a hierarchical framework are not commonplace and hence have been used infrequently. Perhaps more problematic is that such approaches require that study authors fully report the test accuracy results for all grades of dysplasia, and consequently this approach, whilst useful, is not routinely undertaken.

This review revealed some novel cytologic platforms. Several studies (Kämmerer 2013; Kaur 2016; Liu 2019; Ma 2014; Ng 2012) have employed quantitative DNA-image cytometry (DNA-IC). As the complexities of carcinogenesis are revealed through a better understanding of genetic and epigenetic alterations in OPMDs, one might imagine the reality of platforms employing real-time cytology (i.e. chairside results) where morphological parameters are coupled with 'predictive' biomarkers that can identify early lesions with a high risk for malignant transformation. The utility of quantitative cytology to evaluate DNA profile alone and in combination with other markers associated with malignant disease (e.g. keratin 8 and 18 expression) should be undertaken in both primary and secondary care settings (e.g. Ogden 1994). One particularly encouraging approach is the use of point of care brush cytology with high-content single-cell analysis 'cytology on a chip' (Abram 2016; McRae 2020). These well-conducted, large, multicentre case-control test accuracy studies have reported promising test accuracy results at varying levels of positivity threshold. At a positivity threshold of moderate and above, McRae 2020 reported sensitivity of 0.79 (95% confidence interval (CI) 0.74 to 0.83) and specificity of 0.85 (95% CI 0.81 to 0.89), and similarly Abram 2016 reported sensitivity of 0.83 and specificity of 0.72; at a positivity threshold of severe and above, McRae 2020 reported sensitivity of 0.82 (95% CI 0.78 to 0.86) and specificity of 0.88 (95% CI 0.84 to 0.91), and similarly Abram 2016 reported sensitivity of 0.86 and specificity of 0.80. As these studies were case-control test accuracy studies they were not eligible for inclusion in the systematic review, but these important studies should be considered in conjunction with the results for cytology presented in this review.

The diagnostic potential of blood and saliva analysis should also be investigated in future research studies. Despite preliminary work being completed in this area (e.g. Li 2004; Park 2009), study designs have so far been restricted to case-control studies. A recent systematic review highlighted the range of salivary biomarkers under consideration but also the limited amount of evidence per biomarker (AlAli 2020).

ACKNOWLEDGEMENTS

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 Derek Richards for his thoughtful feedback on the protocol; Miguel Ángel González Moles, Rachel Hall, Jennifer Hilgart, Dr Martin G McCabe (The University of Manchester), and the Cochrane Diagnostic Test Accuracy Editorial Team for their feedback on the review; NHS Education Scotland and the Scottish Dental Clinical Effectiveness Programme for their support on this review. We would also like to thank members of the American Dental Association's Center for Evidence Based Dentistry for their continued collaboration and their openness to sharing of data and methodological approaches.

REFERENCES

References to studies included in this review

Adil 2017 *{published data only}*

Adil HA, Yuwanati M, Singh A, Sawant S, Umarji HR. Comparative study on the efficacy of tissue autofluorescence (Visually Enhanced Lesion scope) and toluidine blue as a screening method in oral potentially malignant and malignant lesions. *Journal of Medical Sciences (Taiwan)* 2017;**37**(3):91-6.

Allegra 2009 *{published data only}*

Allegra E, Lombardo N, Puzzo L, Garozzo A. The usefulness of toluidine staining as a diagnostic tool for precancerous and cancerous oropharyngeal and oral cavity lesions [Utilità della colorazione con toluidina nella diagnosi delle lesioni precancerosee cancerose dell'orofaringe e del cavo orale]. *ACTA Otorhinolaryngologica Italica* 2009;**29**:187-90.

Amirchaghmaghi 2018 *{published data only}*

Amirchaghmaghi M, Mohtasham N, Delavarian Z, Shakeri MT, Hatami M, Mosannen Mozafari P. The diagnostic value of the native fluorescence visualization device for early detection of premalignant/malignant lesions of the oral cavity. *Photodiagnosis and Photodynamic Therapy* 2018;**21**:19-27.

Awan 2011 *{published data only}*

* Awan KH, Morgan PR, Warnakulasuriya S. Evaluation of an autofluorescence based imaging system (VELscope™) in the detection of oral potentially malignant disorders and benign keratoses. *Oral Oncology* 2011;**47**:274-7.

Awan KH, Morgan PR, Warnakulasuriya S. Utility of chemiluminescence (ViziLite) in the detection of oral potentially malignant disorders and benign keratoses. *Journal of Oral Pathology & Medicine* 2011;**40**(7):541-4.

Awan 2012 *{published data only}*

Awan KH, Yang YH, Morgan PR, Warnakulasuriya S. Utility of toluidine blue as a diagnostic adjunct in the detection of potentially malignant disorders of the oral cavity – a clinical and histological assessment. *Oral Diseases* 2012;**18**:728-33.

Baeten 2018 *{published data only}*

Baeten J, Johnson A, Sunny S, Suresh A, Birur P, Uma K, et al. Chairside molecular imaging of aberrant glycosylation in subjects with suspicious oral lesions using fluorescently labeled wheat germ agglutinin. *Head & Neck* 2018;**40**(2):292-301.

Cancela-Rodriguez 2011 *{published data only}*

Cancela-Rodriguez P, Cerero-Lapiedra R, Esparza-Gomez G, Llamas-Martinez S, Warnakulasuriya S. The use of toluidine blue in the detection of premalignant and malignant oral lesions. *Journal of Oral Pathology & Medicine* 2010;**40**:300-4.

Chainani-Wu 2015 *{published data only}*

Chainani-Wu N, Madden E, Cox D, Sroussi H, Epstein J, Silverman Jr S. Toluidine blue aids in detection of dysplasia and carcinoma in suspicious oral lesions. *Oral Diseases* 2015;**21**:879-85.

Chaudhari 2013 *{published data only}*

Chaudhari A, Hegde-Shetiya S, Shirahatti R, Agrawal D. Comparison of different screening methods in estimating the prevalence of precancer and cancer amongst male inmates of a jail in Maharashtra, India. *Asian Pacific Journal of Cancer Prevention* 2013;**14**(2):859-64.

Chen 2007 *{published data only}*

Chen Y-W, Lin J-S, Fong JH-J, Wang I-K, Chou S-J, Wu C-H, et al. Use of methylene blue as a diagnostic aid in early detection of oral cancer and precancerous lesions. *British Journal of Oral and Maxillofacial Surgery* 2007;**45**(7):590-1.

Cheng 2003a *{published data only}*

Cheng B, Yang L. The clinical evaluation of Oratest in detecting oral mucosal lesions. *West China Journal of Stomatology* 2003;**21**(2):124-6.

Cheng 2003b *{published data only}*

Cheng B, Yang L. The clinical evaluation of Oratest in detecting oral mucosal lesions. *West China Journal of Stomatology* 2003;**21**(2):124-6.

Chiang 2019 *{published data only}*

Chiang TE, Lin YC, Li YH, Wu CT, Kuo CS, Chen YW. Comparative evaluation of autofluorescence imaging and histopathological investigation for oral potentially malignant disorders in Taiwan. *Clinical Oral Investigations* 2019;**23**(5):2395-402.

Delavarian 2010 *{published data only}*

Delavarian Z, Mohtasham N, Mosannen-Mozaffari P, Pakfetrat A, Shakeri M-T, Ghafoorian-Maddah R. Evaluation of the diagnostic value of a Modified Liquid-Based Cytology using OralCDx Brush in early detection of oral potentially malignant lesions and oral cancer. *Medicina Oral, Patologia Oral y Cirugia Bucal* 2010;**15**(5):e671-6.

Du 2007 *{published data only}*

Du GF, Li CZ, Chen HZ, Chen XM, Xiao Q, Cao ZG, et al. Rose bengal staining in detection of oral precancerous and malignant lesions with colorimetric evaluation: a pilot study. *International Journal of Cancer* 2007;**120**(9):1958-63.

Epstein 2008 *{published data only}*

Epstein JB, Silverman S Jr, Epstein JD, Lonky SA, Bride MA. Analysis of oral lesion biopsies identified and evaluated by visual examination, chemiluminescence and toluidine blue. *Oral Oncology* 2008;**44**:538-44.

Farah 2007 *{published data only}*

Farah CS, McCullough MJ. A pilot case control study on the efficacy of acetic acid wash and chemiluminescent illumination (ViziLite) in the visualisation of oral mucosal white lesions. *Oral Oncology* 2007;**43**:820-4.

Farah 2012 *{published data only}*

Farah CS, McIntosh L, Georgiou A, McCullough MJ. Efficacy of tissue autofluorescence imaging (VELscope) in the visualization of oral mucosal lesions. *Head & Neck* 2012;**32**(6):856-62.

Fontes 2013 *{published data only}*

Fernandes da Costa Fontes KB, Cunha KSG, Rodrigues FR, da Silva LE, Dias EP. Concordance between cytopathology and incisional biopsy in the diagnosis of oral squamous cell carcinoma. *Brazilian Oral Research* 2013;**27**(2):122-7.

Ganga 2017 *{published data only}*

Ganga RS, Gundre D, Bansal S, Shirsat PM, Prasad P, Desai RS. Evaluation of the diagnostic efficacy and spectrum of autofluorescence of benign, dysplastic and malignant lesions of the oral cavity using VELscope. *Oral Oncology* 2017;**75**:67-74.

Guneri 2011 *{published data only}*

Guneri P, Epstein JB, Kaya A, Veral A, Kazandy A, Boyacioglu H. The utility of toluidine blue staining and brush cytology as adjuncts in clinical examination of suspicious oral mucosal lesions. *Oral and Maxillofacial Surgery* 2011;**40**:155-66.

Gupta 2007 *{published data only}*

Gupta A, Singh M, Ibrahim R, Mehrotra R. Utility of toluidine blue staining and brush biopsy in precancerous and cancerous oral lesions. *Acta Cytologica* 2007;**51**(5):788-94.

Hanken 2013 *{published data only}*

Hanken H, Kraatz J, Smeets R, Heiland M, Assaf AT, Blessmann M, et al. The detection of oral premalignant lesions with an autofluorescence based imaging system (VELscope™) - a single-blinded clinical evaluation. *Head & Face Medicine* 2013;**9**:23.

Jayasinghe 2020 *{published data only}*

Jayasinghe RD, Hettiarachchi PVKS, Amugoda D, Kumaraarachchi M, Liyanage RLPR, Siriwardena BSMS, et al. Validity of toluidine blue test as a diagnostic tool for high risk oral potentially malignant disorders - a multicentre study in Sri Lanka. *Journal of Oral Biology and Craniofacial Research* 2020;**10**(4):547-51.

Johnson 2019 *{published data only}*

Johnson A, Baeten J, Patel K, Killian M, Sunny S, Suresh A, et al. Evaluation of a lectin-based imaging system for the chairside detection of oral dysplasia and malignancy. *Journal of Oral and Maxillofacial Surgery* 2019;**77**(9):1941-51.

Kämmerer 2013 *{published data only}*

Kämmerer PW, Koch FP, Santoro M, Babaryka G, Biesterfeld S, Brieger J, et al. Prospective, blinded comparison of cytology and DNA-image cytometry of brush biopsies for early detection of oral malignancy. *Oral Oncology* 2013;**49**(5):420-6.

Kaur 2016 *{published data only}*

Kaur M, Handa U, Mohan H, Dass A. Evaluation of brush cytology and DNA image cytometry for the detection of cancer of the oral cavity. *Diagnostic Cytopathology* 2016;**44**(3):201-5.

Koch 2011a *{published data only}*

Koch FP, Kunkel M, Biesterfeld S. Diagnostic efficiency of differentiating small cancerous and precancerous lesions using mucosal brush smears of the oral cavity - a prospective and blinded study. *Clinical Oral Investigations* 2011;**15**:763-9.

Koch 2011b *{published data only}*

Koch FP, Kaemmerer PW. Effectiveness of autofluorescence to identify suspicious oral lesions - a prospective, blinded clinical trial. *Clinical Oral Investigations* 2011;**15**:975-82.

Leunig 2000 *{published data only}*

Leunig A, Betz CS, Mehlmann M, Stepp H, Arbogast S, Grevers G, et al. Detection of squamous cell carcinoma of the oral cavity by imaging 5-aminolevulinic acid-induced protoporphyrin IX fluorescence. *Laryngoscope* 2000;**110**(1):78-83.

Liu 2019 *{published data only}*

Liu Y, Gao Y, Chen XJ, Hua H. DNA cytometry of exfoliated cells in the diagnosis of oral potential malignant disorders. *Beijing Da Xue Xue Bao [Journal of Peking University. Health Sciences]* 2019;**51**(1):16-20.

Ma 2014 *{published data only}*

Ma J-M, Zhou T-J, Wang R, Shan J, Wu Y-N, Song X-L, et al. Brush biopsy with DNA-image cytometry: a useful and non-invasive method for monitoring malignant transformation of potentially malignant oral disorders. *European Archives of Oto-rhino-laryngology* 2014;**271**(12):3291-5.

Mashberg 1980 *{published data only}*

Mashberg A. Re-evaluation of toluidine blue application as a diagnostic adjunct in the detection of asymptomatic oral squamous carcinoma: a continuing prospective study of oral cancer Ill. *American Cancer Society* 1980;**4**:758-63.

McIntosh 2009 *{published data only}*

McIntosh L, McCullough MJ, Farah CS. The assessment of diffused light illumination and acetic acid rinse (Microlux/DL) in the visualisation of oral mucosal lesions. *Oral Oncology* 2009;**45**:e227-31.

Mehrotra 2008 *{published data only}*

Mehrotra R, Kumar Singh M, Pandya S, Singh M. The use of an oral brush biopsy without computer-assisted analysis in the evaluation of oral lesions: a study of 94 patients. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics* 2008;**106**:246-53.

Mehrotra 2010 *{published data only}*

Mehrotra R, Singh M, Thomas S, Nair P, Pandya S, Shakti Nigam N, et al. A cross-sectional study evaluating chemiluminescence and autofluorescence in the detection of clinically innocuous precancerous and cancerous oral lesions. *Journal of the American Dental Association* 2010;**141**:151-6.

Mehrotra 2011 *{published data only}*

Mehrotra R, Mishra S, Singh M, Singh M. The efficacy of oral brush biopsy with computer assisted analysis in identifying precancerous and cancerous lesions. *Head & Neck Oncology* 2011;**3**(39):1-7.

Mojsa 2012 *{published data only}*

Mojsa I, Kaczmarzyk T, Zaleska M, Stypulkowska J, Zapala-Pospiech A, Sadecki D. Value of the ViziLite Plus System as a diagnostic aid in the early detection of oral cancer/

premalignant epithelial lesions. *Journal of Craniofacial Surgery* 2012;**23**(2):162-4.

Nagaraju 2010 *{published data only}*

Nagaraju K, Prasad S, Ashok L. Diagnostic efficiency of toluidine blue with Lugol's iodine in oral premalignant and malignant lesions. *Indian Journal of Dental Research* 2010;**21**(2):218-23.

Nanayakkara 2016 *{published data only}*

Nanayakkara PG, Dissanayaka WL, Nanayakkara BG, Amaratunga EA, Tilakaratne WM. Comparison of spatula and cytobrush cytological techniques in early detection of oral malignant and premalignant lesions: a prospective and blinded study. *Journal of Oral Pathology & Medicine* 2016;**45**(4):268-74.

Navone 2004 *{published data only}*

Navone R, Marsico A, Reale I, Pich A, Broccoletti R, Pentenero M, et al. Usefulness of oral cytology for the diagnosis of oral squamous dysplasia and carcinoma. *Minerva Stomatologica* 2003;**53**(3):77-86.

Navone 2008 *{published data only}*

Navone R, Pentenero M, Rostan I, Burlo P, Marsico A, Broccoletti R, et al. Oral potentially malignant lesions: first-level micro-histological diagnosis from tissue fragments sampled in liquid-based diagnostic cytology. *Journal of Oral Pathology & Medicine* 2008;**37**:358-63.

Ng 2012 *{published data only}*

Ng SP, Mann IS, Zed C, Doudkine A, Matisic J. The use of quantitative cytology in identifying high-risk oral lesions in community practice. *Oral and Maxillofacial Pathology* 2012;**114**(3):358-64.

Onizawa 1999 *{published data only}*

Onizawa K, Saginoya H, Furuya Y, Yoshida H, Fukuda H. Usefulness of fluorescence photography for diagnosis of oral cancer. *Oral and Maxillofacial Surgery* 1999;**28**:206-10.

Onofre 2001 *{published data only}*

Onofre MA, Sposto MR, Navarro CM. Reliability of toluidine blue application in the detection of oral epithelial dysplasia and in situ and invasive squamous cell carcinomas. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics* 2001;**91**(5):535-40.

Petruzzi 2014 *{published data only}*

Petruzzi M, Lucchese A, Nardi GM, Lauritano D, Favia G, Serpico R, et al. Evaluation of autofluorescence and toluidine blue in the differentiation of oral dysplastic and neoplastic lesions from non-dysplastic and neoplastic lesions: a cross-sectional study. *Journal of Biomedical Optics* 2014;**19**(7):76003.

Rahman 2012 *{published data only}*

Rahman F, Tippu SR, Khandelwal S, Girish KL, Manjunath BC, Bhargava A. A study to evaluate the efficacy of toluidine blue and cytology in detecting oral cancer and dysplastic lesions. *Quintessence International* 2012;**43**:51-9.

Scheer 2011 *{published data only}*

Scheer M, Neugebauer J, Derman A, Fuss J, Drebber U, Zoeller JE. Autofluorescence imaging of potentially malignant mucosa lesions. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics* 2011;**111**(5):568-77.

Scheifele 2004 *{published data only}*

Scheifele C, Schmidt-Westhausen AM, Dietrich T, Reichart PA. The sensitivity and specificity of the OralCDx technique: evaluation of 103 cases. *Oral Oncology* 2004;**40**(8):824-8.

Sciubba 1999 *{published data only}*

Sciubba JJ. Computer-assisted analysis of the oral brush biopsy. *Journal of the American Dental Association* 1999;**130**:1445-57.

Seijas-Naya 2012 *{published data only}*

Seijas-Naya F, García-Carnicero T, Gandara-Vila P, Couso-Folgueiras E, Perez-Sayans M, Gandara-Vila R, et al. Applications of OralCDx methodology in the diagnosis of oral leukoplakia. *Medicina Oral, Patologia Oral y Cirugia Bucal* 2012;**1**(17):5-9.

Sharwani 2006a *{published data only}*

Sharwani A, Jerjes W, Salih V, MacRobert AJ, El-Maaytah M, Khalil HSM, et al. Fluorescence spectroscopy combined with 5-aminolevulinicacid-induced protoporphyrin IX fluorescence in detecting oral premalignancy. *Journal of Photochemistry and Photobiology* 2006;**83**:27-33.

Sharwani 2006b *{published data only}*

Sharwani A, Jerjes W, Salih V, Swinson B, Bigio IJ, El-Maaytah M, et al. Assessment of oral premalignancy using elastic scattering spectroscopy. *Oral Oncology* 2006;**42**:343-9.

Shi 2019 *{published data only}*

Shi L, Li C, Shen X, Zhou Z, Liu W, Tang G. Potential role of autofluorescence imaging in determining biopsy of oral potentially malignant disorders: a large prospective diagnostic study. *Oral Oncology* 2019;**98**:176-9.

Shukla 2018 *{published data only}*

Shukla A, Singh NN, Adsul S, Kumar S, Shukla D, Sood A. Comparative efficacy of chemiluminescence and toluidine blue in the detection of potentially malignant and malignant disorders of the oral cavity. *Journal of Oral and Maxillofacial Pathology* 2018;**22**(3):442.

Silverman 1984 *{published data only}*

Silverman S, Migliorati C, Barbosa J. Toluidine blue staining in the detection of oral precancerous and malignant lesions. *Oral Surgery, Oral Medicine, and Oral Pathology* 1984;**57**(4):379-82.

Singh 2015 *{published data only}*

Singh D, Shukla RK. Utility of toluidine blue test in accessing and detecting intraoral malignancies. *Indian Journal of Otolaryngology and Head & Neck Surgery* 2015;**67 (Suppl 1)**:S47–S50.

Skandarajah 2017 *{published data only}*

Skandarajah A, Sunny SP, Gurpur P, Reber CD, Ambrosio MVD, Raghavan N, et al. Mobile microscopy as a screening

tool for oral cancer in India: a pilot study. *PLOS One* 2017;**12**(11):e0188440.

Svirsky 2002 *{published data only}*

Svirsky JA, Burns JC, Carpenter WM, Cohen DM, Bhattacharyya I, Fantasia JE, et al. Comparison of computer-assisted brush biopsy results with follow-up scalpel biopsy and histology. *General Dentistry* 2002;**50**(6):500-3.

Trakroo 2015 *{published data only}*

Trakroo A, Sunil MK, Trivedi A, Garg R, Kulkarni A, Arora S. Efficacy of oral brush biopsy without computer-assisted analysis in oral premalignant and malignant lesions: a study. *Journal of International Oral Health* 2015;**7**(3):33-8.

Ujaoney 2012 *{published data only}*

Ujaoney S, Motwani MB, Degwekar S, Wadhwan V, Zade P, Chaudhary M, et al. Evaluation of chemiluminescence, toluidine blue and histopathology for detection of high risk oral precancerous lesions: a cross-sectional study. *BMC Clinical Pathology* 2012;**12**(6):1-7.

Upadhyay 2011 *{published data only}*

Upadhyay J, Rao NN, Upadhyay RB, Agarwal P. Reliability of toluidine blue vital staining in detection of potentially malignant oral lesions - time to reconsider. *Asian Pacific Journal of Cancer Prevention* 2011;**12**:1757-60.

Warnakulasuriya 1996 *{published data only}*

Warnakulasuriya KAAS, Johnson NW. Sensitivity and specificity of OraScan toluidine blue mouthrinse in the detection of oral cancer and precancer. *Journal of Oral Pathology and Medicine* 1996;**25**:97-103.

Yamamoto 2017 *{published data only}*

Yamamoto N, Kawaguchi K, Fujihara H, Hasebe M, Kishi Y, Yasukawa M, et al. Detection accuracy for epithelial dysplasia using an objective autofluorescence visualization method based on the luminance ratio. *International Journal of Oral Science* 2017;**9**(11):e2.

References to studies excluded from this review

Abram 2019 *{published data only}*

Abram TJ, Floriano PN, James R, Kerr AR, Thornhill MH, Redding SW, et al. Development of a cytology-based multivariate analytical risk index for oral cancer. *Oral Oncology* 2019;**92**:6-11.

Betz 2002 *{published data only}*

Betz CS, Stepp H, Janda P, Arbogast S, Grevers G, Baumgartner R, et al. A comparative study of normal inspection, autofluorescence and 5-ALA-induced PPIX fluorescence for oral cancer diagnosis. *International Journal of Cancer* 2002;**97**(2):245-52.

Bhoopathi 2009 *{published data only}*

Bhoopathi V, Kabani S, Mascarenhas AK. Low positive predictive value of the oral brush biopsy in detecting dysplastic oral lesions. *Cancer* 2009;**115**(5):1036-40.

Burkhardt 2010 *{published data only}*

Burkhardt A. A response to “sensitivity and specificity of oral brush biopsy” by Hohlweg-Majert et al. *Cancer Investigation* 2010;**28**(5):560-1.

CTRI/2020/01/022999 *{published data only}*

CTRI/2020/01/022999. Usefulness of newly developed light emitting device for screening of precancer and cancer [Validation of novel light emitting device for screening of potential malignant oral disorders and oral cancer]. www.ctri.nic.in/Clinicaltrials/showallp.php?mid1=37933&EncHid=&userName=Usefulness%20of%20newly%20developed%20light%20emitting%20device%20for%20screening%20of%20precancer%20and%20cancer (first received 27 January 2020).

Deuerling 2019 *{published data only}*

Deuerling L, Gaida K, Neumann H, Remmerbach TW. Evaluation of the accuracy of liquid-based oral brush cytology in screening for oral squamous cell carcinoma. *Cancers* 2019;**11**(11):1813.

Driemel 2007 *{published data only}*

Driemel O, Dahse R, Hakim SG, Tsioutsias T, Pistner H, Reichert TE, et al. Laminin-5 immunocytochemistry: a new tool for identifying dysplastic cells in oral brush biopsies. *Cytopathology* 2007;**18**(6):348-55.

Driemel 2007b *{published data only}*

Driemel O, Dahse R, Berndt A, Pistner H, Hakim SG, Zardi L, et al. High-molecular tenascin-C as an indicator of atypical cells in oral brush biopsies. *Clinical Oral Investigations* 2007;**11**:93-9.

Driemel 2008 *{published data only}*

Driemel O, Kunkel M, Hullmann M, Kleinsasser N, Staudenmaier R, Muller-Richter U, et al. Performance of conventional oral brush biopsies [Wertigkeit der konventionellen oralen Bürstenbiopsie]. *HNO* 2008;**56**:205-10.

Ebenezar 2012 *{published data only}*

Ebenezar J, Ganesan S, Aruna P, Muralinaidu R, Renganathan K, Saraswathy TR. Non-invasive fluorescence excitation spectroscopy for the diagnosis of oral neoplasia in vivo. *Journal of Biomedical Optics* 2012;**17**(9):97007.

Epstein 1995 *{published data only}*

Epstein JB, Fatahzadeh M, Matisic J, Anderson G. Exfoliative cytology and electron microscopy in the diagnosis of hairy leukoplakia. *Oral Surgery, Oral Medicine, Oral Pathology* 1995;**79**(5):565-9.

Gomez Serrano 1989 *{published data only}*

Gomez Serrano MT, Fernandez JMT, Bravo JM. Toluidine blue solution as diagnostic aid in oropharyngeal cancer [Azul de toludina en enjuagues, como auxiliar diagnóstico de cancer orofaríngeo]. *Revista ADM* 1989;**46**(1):29-35.

Hedge 2006 *{published data only}*

Hegde MC, Kamath PM, Shreedharan S, Dannana NK, Raju RM. Supravital staining: its role in detecting early malignancies. *Indian Journal of Otolaryngology and Head & Neck Surgery* 2006;**58**(1):31-4.

Hohlweg-Majert 2009 *{published data only}*

Hohlweg-Majert B, Deppe H, Metzger MC, Schumm S, Hoefler H, Kesting MR, et al. Sensitivity and specificity of oral brush biopsy. *Cancer Investigation* 2009;**27**:293-7.

Jayaprakash 2009 *{published data only}*

Jayaprakash V, Sullivan M, Merzianu M, Rigual NR, Loree TR, Popat SR, et al. Autofluorescence-guided surveillance for oral cancer. *Cancer Prevention Research* 2009;**11**(2):966-74.

Kulapaditharon 1998 *{published data only}*

Kulapaditharom B, Boonkitticharroen V. Laser-induced fluorescence imaging in localization of head and neck cancers. *Annals of Otology, Rhinology & Laryngology* 1998;**107**(3):241-6.

Lane 2012 *{published data only}*

Lane P, Lam S, Follen M, MacAulay C. Oral fluorescence imaging using 405-nm excitation, aiding the discrimination of cancers and precancers by identifying changes in collagen and elastic breakdown and neovascularization in the underlying stroma. *Gender Medicine* 2012;**9**(1):S78-82.e8.

Levine 1998 *{published data only}*

Levine TS, Njemenze V, Cowpe JG, Coleman FV. The use of PAPNET automated cytological screening system for the diagnosis of oral squamous carcinoma. *Cytopathology* 1998;**9**(6):398-405.

Li 2004 *{published data only}*

Li Y, St John MA, Zhou X, Kim Y, Sinha U, Jordan RC, et al. Salivary transcriptome diagnostics for oral cancer detection. *Clinical Cancer Research* 2004;**10**(24):8442-50.

Majumder 2006 *{published data only}*

Majumder SK, Gupta A, Gupta S, Ghosh N, Gupta PK. Multi-class classification algorithm for optical diagnosis of oral cancer. *Journal of Photochemistry and Photobiology* 2006;**85**(2):109-17.

Mallia 2010b *{published data only}*

Mallia RJ, Narayanan S, Madhavan J, Sebastian P, Kumar R, Mathews A, et al. Diffuse reflection spectroscopy: an alternative to autofluorescence spectroscopy in tongue cancer detection. *Applied Spectroscopy* 2010;**64**(4):409-13.

Maraki 2004 *{published data only}*

Maraki D, Becker J, Boecking A. Cytologic and DNA-cytometric very early diagnosis of oral cancer. *Journal of Oral Pathology and Medicine* 2004;**33**:398-404.

Maraki 2006 *{published data only}*

Maraki D, Yalcinkaya S, Pomjanski N, Megahed M, Boecking A, Becker J. Cytologic and DNA-cytometric examination of oral lesions in lichen planus. *Journal of Oral Pathology and Medicine* 2006;**35**(4):227-32.

Navone 2007 *{published data only}*

Navone R, Burlo P, Pich A, Pentenero M, Broccoletti R, Marsico A, et al. The impact of liquid-based oral cytology on the diagnosis of oral squamous dysplasia and carcinoma. *Cytopathology* 2007;**18**(6):356-60.

Navone 2009 *{published data only}*

Navone R. Cytology of the oral cavity: a re-evaluation. *Pathologica* 2009;**101**:6-B.

Nieman 2008 *{published data only}*

Nieman LT, Kan CW, Gillenwater A, Markey MK, Sokolov K. Probing local tissue changes in the oral cavity for early detection of cancer using oblique polarized reflectance spectroscopy: a pilot clinical trial. *Journal of Biomedical Optics* 2008;**13**(2):24011-1.

Poate 2004 *{published data only}*

Poate TWJ, Buchanan JAG, Hodgson TA, Speight PM, Barrett AW, Moles DR, et al. An audit of the efficacy of the oral brush biopsy technique in a specialist oral medicine unit. *Oral Oncology* 2004;**40**:829-34.

Rana 2012 *{published data only}*

Ranaa M, Zapfb A, Kuehlea M, Gellrichaand NC, Eckardta AM. Clinical evaluation of an autofluorescence diagnostic device for oral cancer detection: a prospective randomized diagnostic study. *European Journal of Cancer Prevention* 2012;**21**(5):460-6.

Reboires-Lopez 2012 *{published data only}*

Reboiras-López MD, Pérez-Sayáns M, Somoza-Martín JM, Gayoso-Diz P, Barros-Angueira F, Gándara-Rey JM, et al. Comparison of the Cytobrush, dermatological curette and oral CDx brush test as methods for obtaining samples of RNA for molecular analysis of oral cytology. *Cytopathology* 2011;**23**(3):192-7.

Remmerbach 2001 *{published data only}*

Remmerbach TW, Weidenbach H, Pomjanski N, Knops K, Mathes S, Hemprich A, et al. Cytologic and DNA-cytometric early diagnosis of oral cancer. *Analytical Cellular Pathology* 2001;**22**(4):211-21.

Remmerbach 2003 *{published data only}*

Remmerbach TW, Weidenbach H, Müller C, Hemprich A, Pomjanski N, Buckstegge B, et al. Diagnostic value of nucleolar organizer regions (AgNORs) in brush biopsies of suspicious lesions of the oral cavity. *Analytical Cellular Pathology* 2003;**25**(3):139-46.

Remmerbach 2004 *{published data only}*

Remmerbach TW, Mathes SN, Weidenbach H, Hemprich A, Bocking A. Non-invasive brush biopsy as an innovative tool for early detection of oral carcinomas [Nichtinvasive Bürstenbiopsie als innovative Methodein der Früherkennung des Mundhöhlenkarzinoms]. *Mund Kiefer und GesichtsChirurgie* 2004;**8**:229-36.

Remmerbach 2007 *{published data only}*

Remmerbach TW, Hemprich A, Bocking A. [Minimalinvasive Bürstenbiopsie]. *Praxis Fortbildung* 2007;**117**:926-40.

Remmerbach 2009 *{published data only}*

Remmerbach TW, Meyer-Ebrecht D, Aach T, Wurflinger T, Bell AA, Schneider TE, et al. Toward a multimodal cell analysis of brush biopsies for the early detection of oral cancer. *Cancer Cytopathology* 2009;**117**(3):228-35.

Sahu 2017 *{published data only}*

Sahu A, Gera P, Pai V, Dubey A, Tyagi G, Waghmare M, et al. Raman exfoliative cytology for oral precancer diagnosis. *Journal of Biomedical Optics* 2017;**22**(11):1-12.

Sandler 1964 *{published data only}*

Sandler HC. Reliability of oral exfoliative cytology for detection of oral cancer. *Journal of the American Dental Association* 1964;**68**:489-99.

Schwarz 2009 *{published data only}*

Schwarz RA, Gao W, Redden Weber C, Kurachi C, Lee JJ, El-Naggar AK, et al. Non-invasive evaluation of oral lesions using depth-sensitive optical spectroscopy. *Cancer* 2009;**115**(8):1669-79.

Sekine 2017 *{published data only}*

Sekine J, Nakatani E, Hideshima K, Iwahashi T, Sasaki H. Diagnostic accuracy of oral cancer cytology in a pilot study. *Diagnostic Pathology* 2017;**12**(1):27.

Shklar 1970 *{published data only}*

Shklar G, Cataldo E, Meyer I. Reliability of cytologic smear in diagnosis of oral cancer. A controlled study. *Archives of Otolaryngology* 1970;**91**(2):158-60.

Silverman 1992 *{published data only}*

Silverman S, Migliorati C. Toluidine blue staining and early detection of oral precancerous and malignant lesions. *Iowa Dental Journal* 1992;**78**(2):15-6.

Simonato 2017 *{published data only}*

Simonato LE, Tomo S, Miyahara GI, Navarro RS, Villaverde AGJB. Fluorescence visualization efficacy for detecting oral lesions more prone to be dysplastic and potentially malignant disorders: a pilot study. *Photodiagnosis and Photodynamic Therapy* 2017;**17**:1-4.

Sun 2019 *{published data only}*

Sun Z, Liu L, Wang C, Qian L, Yang J, Zhao Z, et al. Assessment of the cancerization risk for oral potentially malignant disorders by clinical risk model combined with autofluorescence and brush biopsy with DNA-image cytometry. *European Archives of Oto-rhino-laryngology* 2019;**276**(9):2549-57.

Swider 1984 *{published data only}*

Swider M, Nakwer K, Wierzbicki J, Grudzinska J. Toluidine test in the diagnosis of oral cancer. *Czasopismo Stomatologiczne* 1984;**37**(9):669-703.

Torres-Rendon 2009 *{published data only}*

Torres-Rendon A, Stewart R, Craig GT, Wells M, Speight PM. DNA ploidy analysis by image cytometry helps to identify oral epithelial dysplasias with a high risk of malignant progression. *Oral Oncology* 2009;**45**:468-73.

Uthoff 2018 *{published data only}*

Uthoff RD, Song B, Sunny S, Patrick S, Suresh A, Kolur T, et al. Point-of-care, smartphone-based, dual-modality, dual-view, oral cancer screening device with neural network classification for low-resource communities. *PLOS One* 2018;**13**(12):e0207493.

Wang 2009 *{published data only}*

Wang CY, Tsai T, Chiang CP, Chen HM, Chen CT. Improved diagnosis of oral premalignant lesions in submucous fibrosis patients with 5-aminolevulinic acid induced PpIX fluorescence. *Journal of Biomedical Optics* 2009;**14**(4):044026.

Yang 2017 *{published data only}*

Yang X, Xiao X, Wu W, Shen X, Zhou Z, Liu W, et al. Cytological study of DNA content and nuclear morphometric analysis for aid in the diagnosis of high-grade dysplasia within oral leukoplakia. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology* 2017;**124**(3):280-5.

References to ongoing studies

CTRI/2018/02/012257 *{unpublished data only}*

CTRI/2018/02/012257. Oral cancer early detection project, a population based study [Early identification of potentially malignant and malignant lesions using clinical oral examination and VELscope on large population]. www.ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=22747&EncHid=&userName=oral%20cancer (first received 28 February 2018).

CTRI/2019/02/017623 *{unpublished data only}*

CTRI/2019/02/017623. Developing an efficient and cost effective method for screening of oral cancer in India [A pilot study to develop an efficacious oral cancer screening strategy for India]. www.ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=29546&EncHid=&userName=oral%20cancer (first received 12 February 2019).

CTRI/2019/11/022167 *{published data only}*

CTRI/2019/11/022167. Mobile oral cancer screening in limited resource setting [Low-cost mobile oral cancer screening for low resource setting]. www.ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=37024&EncHid=&userName=oral%20cancer (first received 27 November 2019).

Additional references

Abram 2016

Abram TJ, Floriano PN, Christodoulides N, James R, Kerr AR, Thornhill MH, et al. 'Cytology-on-a-chip' based sensors for monitoring of potentially malignant oral lesions. *Oral Oncology* 2016;**60**:103-11.

AlAli 2020

AlAli AM, Walsh T, Maranzano M. CYFRA 21-1 and MMP-9 as salivary biomarkers for the detection of oral squamous cell carcinoma: a systematic review of diagnostic test accuracy. *International Journal of Oral and Maxillofacial Surgery* 2020;**49**(8):973-83.

Arduino 2013

Arduino PG, Bagan J, El-Naggar AK, Carrozzo M. Urban legends series: oral leukoplakia. *Oral Diseases* 2013;**19**(7):642-59.

Bagan 2020

Bagan JV, Bagan-Debon L. Clinical presentation and differential diagnosis of oral cancer. In: Warnakulasuriya S, Greenspan JS, editors(s). Textbook of Oral Cancer. Switzerland: Springer Nature, 2020:47-54.

Bossuyt 2003

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

Bravi 2021

Bravi F, Lee YA, Hashibe M, Boffetta P, Conway DI, Ferraroni M, et al. Lessons learned from the INHANCE consortium: an overview of recent results on head and neck cancer. *Oral Diseases* 2021;**27**(1):73-93.

Bray 2018

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a Cancer Journal for Clinicians* 2018;**68**(6):394–424.

Brinkmann 2011

Brinkmann O, Kastratovic DA, Dimitrijevic MV, Konstantinovic VS, Jelovac DB, Antic J, et al. Oral squamous cell carcinoma detection by salivary biomarkers in a Serbian population. *Oral Oncology* 2011;**47**(1):51-5.

Brocklehurst 2013

Brocklehurst P, Kujan O, O'Malley LA, Ogden G, Shepherd S, Glenny AM. Screening programmes for the early detection and prevention of oral cancer. *Cochrane Database of Systematic Reviews* 2013, Issue 11. Art. No: CD004150. [DOI: 10.1002/14651858.CD004150.pub3]

Buchen 2011

Buchen L. Cancer: missing the mark. *Nature* 2011;**471**(7339):428-32.

Bulsara 2018

Bulsara VM, Worthington HV, Glenny AM, Clarkson JE, Conway DI, Macluskey M. Interventions for the treatment of oral and oropharyngeal cancers: surgical treatment. *Cochrane Database of Systematic Reviews* 2018, Issue 12. Art. No: CD006205. [DOI: 10.1002/14651858.CD006205.pub4]

Cancer Research UK

Cancer Research UK. Oral cancer statistics. www.cancerresearchuk.org/cancer-info/cancerstats/types/oral (accessed 8 April 2011).

Chaturvedi 2019

Chaturvedi AK, Udaltsova N, Engels EA, Katzel JA, Yanik EL, Katki HA, et al. Oral leukoplakia and risk of progression to oral cancer: a population-based cohort study. *Journal of the National Cancer Institute* 2020;**112**(10):1047-54.

Chi 2015

Chi AC, Day TA, Neville BW. Oral cavity and oropharyngeal squamous cell carcinoma - an update. *CA: a Cancer Journal for Clinicians* 2015;**65**(5):401-21.

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.

de Pauli Paglioni 2020

de Pauli Paglioni M, Migliorati CA, Schausltz Pereira Faustino I, Linhares Almeida Mariz BA, Oliveira Corrêa Roza AL, Agustin Vargas P, et al. Laser excision of oral leukoplakia: does it affect recurrence and malignant transformation? A systematic review and meta-analysis. *Oral Oncology* 2020;**109**:104850. [DOI: doi:10.1016/j.oraloncology.2020.104850]

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.

Divani 2009

Divani S, Exarhou M, Theodorou LN, Georgantzis D, Skoulakis H. Advantages and difficulties of brush cytology in the identification of early oral cancer. *Archive of Oncology* 2009;**17**(1-2):11-2.

Dost 2014

Dost F, Lê Cao K, Ford PJ, Ades C, Farah CS. Malignant transformation of oral epithelial dysplasia: a real-world evaluation of histopathologic grading. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology* 2014;**117**(3):343-52.

ERO-FDI 2019

Perks A, Forna DA, Barreira E, Fricain JC, Monteiro L, Diniz Freitas M, et al. Oral Cancer Prevention. European Regional Organization of the Federation Dentaire Internationale (FDI), 2019.

Farah 2019

Farah CS, Shearston K, Nguyen AP, Kujan O. Oral carcinogenesis and malignant transformation. In: Brennan PA, Aldridge T, Dwivedi RC, editors(s). Premalignant Conditions of the Oral Cavity. Singapore: Springer Singapore, 2019:27-66.

Fedele 2009

Fedele S. Diagnostic aids in the screening of oral cancer. *Head and Neck Oncology* 2009;**1**:5.

Ferlay 2010

Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer* 2010;**127**(12):2893-917.

Fitzpatrick 2014

Fitzpatrick SG, Hirsch SA, Gordon SC. The malignant transformation of oral lichen planus and oral lichenoid lesions: a systematic review. *Journal of the American Dental Association* 2014;**145**(1):45-56.

Furness 2011

Furness S, Glenny AM, Worthington HV, Pavitt S, Oliver R, Clarkson JE, et al. Interventions for the treatment of oral cavity and oropharyngeal cancer: chemotherapy. *Cochrane Database of Systematic Reviews* 2011, Issue 4. Art. No: CD006386. [DOI: 10.1002/14651858.CD006386.pub3]

Ganly 2012

Ganly I, Patel S, Shah J. Early stage squamous cell cancer of the oral tongue - clinicopathologic features affecting outcome. *Cancer* 2012;**118**(1):101-11.

Garg 2005

Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005;**293**(10):1223-38.

Giuliani 2019

Giuliani M, Troiano G, Cordaro M, Corsalini M, Gioco G, Lo Muzio L, et al. Rate of malignant transformation of oral lichen planus: a systematic review. *Oral Diseases* 2019;**25**(3):693-709.

Glenny 2010

Glenny AM, Furness S, Worthington HV, Conway DI, Oliver R, Clarkson JE, et al. Interventions for the treatment of oral cavity and oropharyngeal cancer: radiotherapy. *Cochrane Database of Systematic Reviews* 2010, Issue 12. Art. No: CD006387. [DOI: 10.1002/14651858.CD006387.pub2]

González-Moles 2019

González-Moles MÁ, Ruiz-Ávila I, González-Ruiz L, Ayén Á, Gil-Montoya JA, Ramos-García P. Malignant transformation risk of oral lichen planus: a systematic review and comprehensive meta-analysis. *Oral Oncology* 2019;**86**:121-30.

Harbord 2009

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

Holmstrup 2006

Holmstrup P, Vedtofte P, Reibel J, Stoltze K. Long-term treatment outcome of oral premalignant lesions. *Oral Oncology* 2006;**42**(5):461-74.

IARC 2012

World Health Organization, International Agency for Research on Cancer (IARC). Personal habits and indoor combustions. Lyon (France): IARC Monographs on the Evaluation of Carcinogenic Risks to Humans; 2012. Volume 100E; I2012.

Idrees 2021

Idrees M, Kujan O, Shearston K, Farah CS. Oral lichen planus has a very low malignant transformation rate: a systematic review and meta-analysis using strict diagnostic and inclusion criteria. *Journal of Oral Pathology & Medicine* 2021;**50**(3):287-98.

Jin 2016

Jin LJ, Lamster IB, Greenspan JS, Pitts NB, Scully C, Warnakulasuriya S. Global burden of oral diseases: emerging concepts, management and interplay with systemic health. *Oral Diseases* 2016;**22**(7):609-19.

Kerr 2020

Kerr AR. Diagnostic adjuncts for oral cavity squamous cell carcinoma and oral potentially malignant disorders. In: Warnakulasuriya S, Greenspan JS, editors(s). Textbook of Oral Cancer. Switzerland: Springer Nature, 2020:99-118.

Kreimer 2020

Kreimer AR, Chaturvedi AK, Alemany L, Anantharaman D, Bray F, Carrington M, et al. Summary from an international cancer seminar focused on human papillomavirus (HPV)-positive oropharynx cancer, convened by scientists at IARC and NCI. *Oral Oncology* 2020;**108**:104736.

Landis 1977

Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**(1):159-74.

Lee 2009

Lee JM, Garon E, Wong DT. Salivary diagnostics. *Orthodontics and Craniofacial Research* 2009;**12**(3):206-11.

Leeflang 2008

Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine* 2008;**149**(12):889-97.

Leston 2010

Seoane Lestón J, Diz Dios P. Diagnostic clinical aids in oral cancer. *Oral Oncology* 2010;**46**(6):418-22.

Li 2006

Li Y, Elashoff D, Oh M, Sinha U, St John MA, Zhou X, et al. Serum circulating human mRNA profiling and its utility for oral cancer detection. *Journal of Clinical Oncology* 2006;**24**(11):1754-60.

Lim 2003

Lim K, Moles DR, Speight PM. Opportunistic screening for oral cancer and precancer in general dental practice: results of a demonstration study. *British Dental Journal* 2003;**194**:497–502.

Lingen 2008

Lingen MW, Kalmar JR, Karrison T, Speight PM. Critical evaluation of diagnostic aids for the detection of oral cancer. *Oral Oncology* 2008;**44**(1):10-22.

Lingen 2017

Lingen MW, Tampi MP, Urquhart O, Abt E, Agrawal N, Chaturvedi AK, et al. Adjuncts for the evaluation of potentially

malignant disorders in the oral cavity: diagnostic test accuracy systematic review and meta-analysis - a report of the American Dental Association. *Journal of the American Dental Association* 2017;**148**(11):797-813.

Liu 2006

Liu JL, Wyatt JC, Deeks JJ, Clamp S, Keen J, Verde P, et al. Systematic reviews of clinical decision tools for acute abdominal pain. *Health Technology Assessment* 2006;**10**(47):1-167.

Liu 2016

Liu D, Zhao X, Zeng X, Dan H, Chen Q. Non-invasive techniques for detection and diagnosis of oral potentially malignant disorders. *Tohoku Journal of Experimental Medicine* 2016;**238**(2):165-77.

Locca 2020

Locca O, Sollecito TP, Alawi F, Weinstein GS, Newman JG, De Virgilio A, et al. Potentially malignant disorders of the oral cavity and oral dysplasia: a systematic review and meta-analysis of malignant transformation rate by subtype. *Head & Neck* 2020;**42**(3):539-55.

Lodi 2006

Lodi G, Sardella A, Bez C, Demarosi F, Carrassi A. Interventions for treating oral leukoplakia. *Cochrane Database of Systematic Reviews* 2006, Issue 4. Art. No: CD001829. [DOI: 10.1002/14651858.CD001829.pub3]

Lodi 2016

Lodi G, Franchini R, Warnakulasuriya S, Varoni EM, Sardella A, Kerr AR, et al. Interventions for treating oral leukoplakia to prevent oral cancer. *Cochrane Database of Systematic Reviews* 2016, Issue 7. Art. No: CD001829. [DOI: 10.1002/14651858.CD001829.pub4]

Macaskill 2011

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, 2011. Available from methods.cochrane.org/sdt/handbook-dta-reviews.

Madhura 2020

Madhura MG, Rao RS, Patil S, Fageeh HN, Alhazmi A, Awan KH. Advanced diagnostic aids for oral cancer. *Disease-a-month* 2020;**66**(12):101034. [DOI: 10.1016/j.disamonth.2020.101034]

McRae 2020

McRae MP, Modak SS, Simmons GW, Trochesset DA, Kerr AR, Thornhill MH, et al. Point-of-care oral cytology tool for the screening and assessment of potentially malignant oral lesions. *Cancer Cytopathology* 2020;**128**(3):207-20.

Mehanna 2009

Mehanna HM, Rattay T, Smith J, McConkey CC. Treatment and follow-up of oral dysplasia - a systematic review and meta-analysis. *Head & Neck* 2009;**31**(12):1600-9.

Miranda-Filho 2020

Miranda-Filho A, Bray F. Global patterns and trends in cancers of the lip, tongue and mouth. *Oral Oncology* 2020;**102**:104551. [DOI: 10.1016/j.oraloncology.2019.104551]

NICE 2016

National Institute for Health and Care Excellence (NICE). Head and neck cancers - recognition and referral. cks.nice.org.uk/topics/head-neck-cancers-recognition-referral/ 2016.

Nikitakis 2018

Nikitakis NG, Pentenero M, Georgaki M, Poh CF, Peterson DE, Edwards P, et al. Molecular markers associated with development and progression of potentially premalignant oral epithelial lesions: current knowledge and future implications. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology* 2018;**125**(6):650-69.

Ogden 1994

Ogden GR, Cowpe JG, Chisholm DM, Lane EB. DNA and keratin analysis of oral exfoliative cytology in the detection of oral cancer. *Oral Oncology, European Journal of Cancer* 1994;**30B**:405-8.

Ogden 2015

Ogden GR, Scully C, Warnakulasuriya S, Speight P. Oral cancer: two cancer cases in a career? *British Dental Journal* 2015;**218**:439.

Ogden 2020

Ogden GR. Factors affecting survival for oral cancer. In: Warnakulasuriya S, Greenspan JS, editors(s). Textbook of Oral Cancer. Switzerland: Springer Nature, 2020:327-42.

Omar 2015

Omar E. Current concepts and future of non-invasive procedures for diagnosing oral squamous cell carcinoma - a systematic review. *Head & Face Medicine* 2015;**11**:6.

Park 2009

Park NJ, Zhou H, Elashoff D, Henson BS, Kastratovic DA, Abemayor E, et al. Salivary microRNA: discovery, characterization, and clinical utility for oral cancer detection. *Clinical Cancer Research* 2009;**15**(17):5473-7.

Parkin 2001

Parkin DM. Global cancer statistics in the year 2000. *Lancet Oncology* 2001;**2**(9):533-43.

Patton 2008

Patton LL, Epstein JB, Kerr AR. Adjunctive techniques for oral cancer examination and lesion diagnosis: a systematic review of the literature. *Journal of the American Dental Association* 2008;**139**(7):896-905.

Petti 2003

Petti S. Pooled estimate of world leukoplakia prevalence: a systematic review. *Oral Oncology* 2003;**39**(8):770-80.

Purkayastha 2018

Purkayastha M, McMahon A, Gibson J, Conway DI. Is detecting oral cancer in general dental practices a realistic expectation? A population-based study using population linked data in Scotland. *British Dental Journal* 2018;**225**:241-6.

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982-90.

Rethman 2010

Rethman MP, Carpenter W, Cohen EE, Epstein J, Evans CA, Flaitz CM, et al. Evidence-based clinical recommendations regarding screening for oral squamous cell carcinomas. *Journal of the American Dental Association* 2010;**141**(5):509-20.

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.

Rusthoven 2010

Rusthoven KE, Raben D, Song JI, Kane M, Altoos TA, Chen C. Survival and patterns of relapse in patients with oral tongue cancer. *Journal of Oral and Maxillofacial Surgery* 2010;**68**(3):584-9.

Schünemann 2020

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

Schünemann 2020a

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

Seoane 2016

Seoane J, Alvarez-Novoa P, Gomez I, Takkouche B, Diz P, Warnakulasiruya S, et al. Early oral cancer diagnosis: the Aarhus statement perspective. A systematic review and meta-analysis. *Head & Neck* 2016;**38 Suppl 1**:E2182-9.

Shaw 2020

Shaw RJ, O'Connell JE, Bajwa M. Basic surgical principles and techniques. In: Warnakulasuriya S, Greenspan JS, editors(s). Textbook of Oral Cancer. Switzerland: Springer Nature, 2020:253-82.

Shi 2019

Shi L, Li C, Shen X, Zhou Z, Liu W, Tang G. Potential role of autofluorescence imaging in determining biopsy of oral potentially malignant disorders: a large prospective diagnostic study. *Oral Oncology* 2019;**98**:176-9.

Shield 2017

Shield KD, Ferlay J, Jemal A, Sankaranarayanan R, Chaturvedi AK, Bray F, et al. The global incidence of lip, oral cavity, and pharyngeal cancers by subsite in 2012. *CA: a Cancer Journal for Clinicians* 2017;**67**(1):51-6.

Speight 2017

Speight PM, Epstein J, Kujan O, Lingen MW, Nagao T, Ranganathan K, et al. Screening for oral cancer - a perspective from the Global Oral Cancer Forum. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology* 2017;**123**(6):680-7.

Takwoingi 2016

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

Tang 2000

Tang JL, Liu JL. Misleading funnel plot for detection of bias in meta-analysis. *Journal of Clinical Epidemiology* 2000;**53**(5):477-84.

Thompson 2017

Thomson PJ, Goodson ML, Smith DR. Profiling cancer risk in oral potentially malignant disorders - a patient cohort study. *Journal of Oral Pathology & Medicine* 2017;**46**(10):888-95.

Tiwari 2019

Tiwari L, Kujan O, Farah CS. Optical fluorescence imaging in oral cancer and potentially malignant disorders: a systematic review. *Oral Diseases* 2020;**26**(3):491-510.

Walsh 2013

Walsh T, Liu JLY, Brocklehurst P, Glenny AM, Lingen M, Kerr AR, et al. Clinical assessment to screen for the detection of oral cavity cancer and potentially malignant disorders in apparently healthy adults. *Cochrane Database of Systematic Reviews* 2013, Issue 11. Art. No: CD010173. [DOI: 10.1002/14651858.CD010173.pub2]

Warnakulasuriya 2007

Warnakulasuriya S, Johnson NW, van der Waal I. Nomenclature and classification of potentially malignant disorders of the oral mucosa. *Journal of Oral Pathology and Medicine* 2007;**36**(10):575-80.

Warnakulasuriya 2009

Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. *Oral Oncology* 2009;**45**(4-5):309-16.

Warnakulasuriya 2011

Warnakulasuriya S, Kovacevic T, Madden P, Coupland VH, Sperandio M, Odell E, et al. Factors predicting malignant transformation in oral potentially malignant disorders among patients accrued over a 10-year period in South East England. *Journal of Oral Pathology and Medicine* 2011;**40**(9):677-83.

Warnakulasuriya 2016

Warnakulasuriya S, Ariyawardana A. Malignant transformation of oral leukoplakia: a systematic review of observational

studies. *Journal of Oral Pathology and Medicine* 2016;**45**(3):155-6.

Warnakulasuriya 2020

Warnakulasuriya S, Greenspan JS. Epidemiology of oral and oropharyngeal cancers. In: Warnakulasuriya S, Greenspan JS, editors(s). Textbook of Oral Cancer. Springer Nature: Switzerland, 2020:5-22.

Whiting 2011

Whiting PF, Rutjes AW, 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**(8):529-36.

Wyatt 1995

Wyatt JC, Altman DG. Commentary: prognostic models: clinically useful or quickly forgotten? *BMJ* 1995;**311**(7019):1539-41.

CHARACTERISTICS OF STUDIES

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

Adil 2017

Study characteristics	
Patient Sampling	Method of patient selection: "ninety patients with tobacco-associated hyperkeratotic red and white lesion, ulcerative lesion, and frank malignancy. Suspicious-looking red and white lesions without any history of tobacco were also included in this study. Patients who presented with hypermelanotic lesion associated with tobacco habit were excluded"
Patient characteristics and setting	Age: 22 to 70 years Sex: 15 female, 75 male SES: not reported Ethnicity: not reported Stated risk factors: 41% smokers, alcohol consumption not reported Number of patients/lesions: 90/90 Lesion site: buccal mucosa 70 cases, tongue 7, labial mucosa 6, palatal mucosa 4, and floor of mouth 1 Severity: "Out of 56 potentially malignant lesions, 34 were speckled leukoplakia, 18 were homogeneous leukoplakia, 3 were verrucous leukoplakia, and 1 was erythroplakia" Country: India Type of facility: secondary care facility assumed Prevalence: 82/90
Index tests	Category: light-based - VELscope, staining - toluidine blue Description: quote: "the lesion was subjected to standard 1% toluidine blue staining" Positivity threshold: quote: "Greenish fluorescence emanating from site was considered as indicative of normal and healthy mucosa. Dark/brownish color fluorescence observed through VELscope was considered as positive for dysplastic changes within the mucosa." Quote: "Dark blue (royal or navy) staining of lesion was considered positive for dysplasia [Figure 3]. No staining or light blue staining of lesion was considered as negative for dysplasia"

References to other published versions of this review

Liu 2012

Liu JLY, Walsh T, Kerr AR, Lingen M, Brocklehurst P, Ogden G, et al. Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions. *Cochrane Database of Systematic Reviews* 2012, Issue 12. Art. No: CD010276. [DOI: 10.1002/14651858.CD010276]

Macey 2015

Macey R, Walsh T, Brocklehurst P, Kerr AR, Liu JLY, Lingen MW, et al. Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions. *Cochrane Database of Systematic Reviews* 2015, Issue 5. Art. No: CD010276. [DOI: 10.1002/14651858.CD010276.pub2]

* Indicates the major publication for the study

Adil 2017 *(Continued)*

	Sequence of tests: index (VELscope, toluidine blue) followed by reference Training or calibration of clinicians: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: based on clinical diagnosis Conflict of interests: authors declare no conflict of interest
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: incisional biopsy Positivity threshold: quote: "oral potentially malignant lesions or malignant lesions" Sequence of tests: index then reference Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: quotes: "After scrutinizing photographic evidence of VELscope finding and toluidine blue testing, sites for surgical biopsy were decided" "biopsy sites were selected based on fluorescence loss and/or toluidine blue stain test. In case both the VELscope and toluidine blue test were negative, the site for biopsy was dictated by clinical examination" Target condition: quote: "The dysplastic lesions were graded according to the number of dysplastic features exhibited into mild, moderate, and severe dysplasia"
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported but assumed following index tests Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Results from the toluidine blue test could possibly be excluded from the meta-analysis in a sensitivity analysis. The 2 tests may not be considered to be independent and so only the first test (VELscope) could be used. However, nothing is explicitly stated in the manuscript to say that lesions identified by VELscope were subsequently given toluidine blue staining

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Adil 2017 *(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		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Allegra 2009

Study characteristics	
Patient Sampling	Method of patient selection: quote: "The study focuses on 45 oral mucosa lesions from 32 patients (13 female, 19 male, mean age 59 years, range 42-82), coming under observation at the Department of Otolaryngology - Head and Neck Surgery, at the University of Catanzaro"
Patient characteristics and setting	Age: mean age 59 years (range 42 to 82) Sex: 13 female, 19 male SES: not reported Ethnicity: not reported Stated risk factors: none stated Number of patients/lesions: 32/45 Lesion site: quote: "The site of the lesion was tongue in 11 cases, buccal mucosa in 9, floor of the mouth in 8 and hard or soft palate in 4"

Allegra 2009 *(Continued)*

	Severity: non-neoplastic (hyper-keratoses, hyper-para-keratoses, etc.), mild dysplasia, moderate dysplasia, severe dysplasia, in situ carcinoma, invasive carcinoma Country: Italy Type of facility: Department of Otolaryngology - Head and Neck Surgery, at the University of Catanzaro Prevalence: 30/45 lesions were premalignant or malignant
Index tests	Category: vital staining - toluidine blue Description: quote: "Patients rinsed the oral cavity with water for 20 sec. to remove debris prior to rinsing with 1% acetic acid for 20 sec. Toluidine blue (1% W/W) was applied as an oral rinse for 20 sec. and then 1% acetic acid was used for 20 sec to eliminate mechanically retained stain" Positivity threshold: quote: "Lesions that showed dark blue staining were considered to be positive for premalignant or malignant tissue, while those with light staining, or totally not coloured, were considered negative" Sequence of tests: staining followed by reference standard Training or calibration of clinicians: no training reported Blinding of examiners: index test completed before reference standard Multiple tests: no Method of site selection: toluidine blue rinse - no site selection Conflict of interests: not stated
Target condition and reference standard(s)	Category: biopsy (punch) with histopathologic assessment Description: quote: "The biopsies were performed under local anaesthesia by punch biopsy, all specimens were labelled with a progressive number and in a separate book, for each specimen the clinical examination and the result of the toluidine blue staining were reported" Positivity threshold: quote: "Histopathologic diagnoses were referred as: non-neoplastic (hyper-keratoses, hyper-para-keratoses, etc.), mild dysplasia, moderate dysplasia, severe dysplasia, in situ carcinoma, invasive carcinoma" Sequence of tests: index test followed by reference standard Training or calibration of pathologists: not stated Blinding of examiners: index test completed before reference standard. Quote: "The pathologist examining all the biopsies was not informed regarding the clinical or staining evaluation of each sample" Multiple tests: no Method of site selection: unclear Target condition: precancerous and cancerous oropharyngeal and oral cavity lesions
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: within same appointment so minimal interval Patients receiving both index and reference test but excluded from analysis: 0

Allegra 2009 *(Continued)*

Comparative
User defined 1
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		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Allegra 2009 *(Continued)*

Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Amirchaghmaghi 2018

Study characteristics	
Patient Sampling	Method of patient selection: consecutive sampling suggested: "All patients over 18 years of age with oral soft tissue lesions who required incisional or excisional biopsy for further diagnosis were included in the study" Exclusions: quote: "Patients with contraindications for biopsy sampling, such as hemorrhagic diseases or uncontrolled systemic diseases, or patients with a confirmed diagnosis of dysplasia or malignancy in a previous biopsy were excluded from the study"
Patient characteristics and setting	Age: 52.3 years (SD 14.8) Sex: 24 female, 21 male SES: not reported Ethnicity: not reported Stated risk factors: tobacco use recorded Number of patients/lesions: 50 patients invited but only 45 eligible, 54 lesions Lesion site: oral cavity Severity: benign, dysplastic, and malignant oral lesions Country: Iran Type of facility: "Oral and Maxillofacial Medicine Department of the School of Dentistry, Mashhad University" Prevalence: 21/54 lesions with dysplasia or malignant carcinoma
Index tests	Category: light-based - VELscope Description: conventional oral cavity examination was performed prior to the fluorescence visualisation of the oral cavity. Quote: "the location and approximate loss of fluorescence areas and regions irradiating red/orange light were drawn as a schematic illustration with different colors" Positivity threshold: quote: "Regions with LOF or that were seen as red/orange were considered to be suspicious sites (positive VEL)" Sequence of tests: index followed by reference standard Training or calibration of clinicians: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: any lesion in the oral cavity Conflict of interests: authors declare that they have no conflicts of interest

Amirchaghmaghi 2018 *(Continued)*

Target condition and reference standard(s)	Category: incisional biopsy and histopathology Description: biopsies taken from sites identified by conventional examination and VELscope Positivity threshold: with dysplasia or malignant carcinoma Sequence of tests: followed index test Training or calibration of pathologists: quote: "experienced specialist in oral and maxillofacial pathology" Blinding of examiners: not reported Multiple tests: no Method of site selection: quotes: "the selection of biopsy sites were done after the consensus of two specialists in oral and maxillofacial medicine" "For premalignant lesions/conditions without any evidence of dysplasia (COE negative), the biopsy sample was taken from locations that are surgically accessible" Target condition: mild moderate and severe dysplasia or malignant carcinoma
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported, no reason to expect a time interval between tests Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	The reported results used in the meta-analysis are those in which the specialist conventional examination and the fluorescence have been combined

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Amirchaghmaghi 2018 *(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		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	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?		Low risk	

Awan 2011

Study characteristics	
Patient Sampling	Method of patient selection: quote: "One hundred and sixty-four consecutive patients aged over 16 years presenting in oral medicine clinics at two London hospitals with white, red, and mixed white and red patches were invited to participate in the study. One hundred and twenty-six patients (76.8%) consented and were investigated by a standard protocol that involved clinical visual examination and autofluorescence examination followed by biopsy." Of the 126 patients 116 received the reference test
Patient characteristics and setting	Setting: secondary care, oral medicine clinics Age: over 16 years Sex: male 70 (55.6%), female 56 (44.4%) SES: not reported Ethnicity: white 76 (60.3%), non-white 50 (39.7%) Stated risk factors: smokers 61 (48.4%); ex-smoker 28 (22.2%); never smoked 37 (29.4%)

Awan 2011 *(Continued)*

alcohol: current user 92 (73%); ex-user 8 (6.4%); never used 26 (20.6%)

Number of patients/lesions: 126/126, 10 did not receive reference test

Lesion site: buccal mucosa 54 (42.9%), tongue 40 (31.7%), floor of mouth 14 (11.1%), palate 11 (8.7%), alveolar ridge 7 (5.5%)

Severity: unclear prior to testing, described as "with white, red, and mixed white and red patches"

Country: UK

Type of facility: oral medicine clinic at 2 hospitals

Prevalence: dysplastic: 44/116 (VELscope and ViziLite)

Index tests

Index test 1 - VELscope

Category: light-based

Description: the main lesion was selected and diagnosed by consensus of 2 experienced examiners after comprehensive clinical examination, photographed then examined with the VELscope; quote: "under dimmed room light, with protective eye wear worn by the patient throughout the procedure"

Positivity threshold: quote: "The possible outcome of the autofluorescence examination was determined by the manufacturer's literature i.e. FVL – fluorescence visualization loss, FVR – fluorescence visualization retained and FVI – fluorescence visualization increased"

Sequence of tests: VELscope then ViziLite then toluidine blue, followed by reference standard

Training or calibration of clinicians: calibrated by an experienced professional from the manufacturer, no results reported

Blinding of examiners: index test examiners blinded to reference test results

Multiple tests: yes; conducted independently during the same session, by the same examiners. Order: VELscope, ViziLite, toluidine blue

Method of site selection: principle area of morphology decided by consensus after visual examination

Conflict of interests: equipment provided by LED Diagnosis who manufacture and market the VELscope

Index test 2 - ViziLite

Category: light-based

Description: chemiluminescent light source adjunct to oral examination; quote: "The oral cavity was rinsed with 1% acetic acid solution prior to examination with the ViziLite"

Positivity threshold: quote: "lesions that showed an aceto-white appearance under the chemiluminescent light"

Sequence of tests: VELscope then ViziLite then toluidine blue, followed by reference standard

Training or calibration of clinicians: 2 clinicians, 1 of whom was an experienced examiner. No reporting of training or calibration by demonstration of kappa score, although does report that COE findings, the area for further examination identified, and findings using ViziLite were all verified and consensus found between the 2 clinicians

Blinding of examiners: index test examiners blinded to reference test results

Multiple tests: yes; conducted independently during the same session, by the same examiners. Order: VELscope, ViziLite, toluidine blue

Method of site selection: quote: "The principal area (site) of morphologically altered mucosa excluding any ulcerated areas (by consensus of both examiners) was selected and photographed. All further investigations were performed on this clinically detected area of mucosal abnormality"

Awan 2011 *(Continued)*

	Conflict of interests: equipment provided by manufacturers and marketers of ViziLite; quote: "ViziLite kits for the study were supplied by Zila Inc."
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "A surgical biopsy was performed for histopathological assessment... the presence or absence of dysplasia was further reviewed by an experienced oral pathologist" Positivity threshold: quote: "The presence or absence of dysplasia in the biopsy specimen was recorded by an experienced oral pathologist" Sequence of tests: index tests followed by reference standard Training or calibration of pathologists: only 1 experienced pathologist Blinding of examiners: quote: "A surgical biopsy of the clinically altered area was performed for histopathological assessment, and after formal diagnostic reporting by two pathologists (blinded to ViziLite data), the presence or absence of dysplasia was further reviewed by an experienced oral pathologist." Each index test reported separately, quote taken from ViziLite paper, but not clarified in other paper Multiple tests: no Method of site selection: the index test results were used to inform the selection of biopsy site Target condition: dysplasia
Flow and timing	Patients receiving index test but not reference test: 10 Patients receiving reference test but not index test: 0 Time interval: the tests were conducted independently (though on the same session for patient convenience) Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	**Index test - VELscope & ViziLite:** only able to report on 116 patients that received reference test. Papers report data on leukoplakia and erythroplakia (through clinical diagnosis only), only data reporting dysplastic or non-dysplasia from pathology used Results published for ViziLite are misprinted in the journal article, 72 classed as "other", made up of 52 aceto-white and 20 normal. Confirmed by author

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients		Low risk	

Awan 2011 *(Continued)*

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

Awan 2012

Study characteristics	
Patient Sampling	Method of patient selection: consecutive patients presenting in oral medicine clinics at 2 London hospitals, with white, red, and mixed white and red patches
Patient characteristics and setting	Setting: secondary care, oral medicine clinics Age: over 16 years Sex: male 61%, female 39% SES: not reported Ethnicity: white 67%, non-white 33% Stated risk factors: smokers 52%; ex-smoker 22%; never smoked 26% alcohol: current user 79%; ex-user 6%; never used 15% Number of patients/lesions: 92/92 Lesion site: buccal mucosa 37%, tongue 33%, floor of mouth 11%, palate 11%, alveolar ridge 9% Severity: unclear prior to testing, described as "with white, red, and mixed white and red patches" Country: UK Type of facility: oral medicine clinic at 2 hospitals Prevalence: 41/92
Index tests	Category: vital staining Description: quote: "TBlue staining test was performed using the TBlue oral lesion marking system.... The TBlue kit consisted of three swab tubes: Swab tube1, 1% acetic acid solution (pre rinse swab), Swab tube 2, 0.5% tolonium chloride solution and Swab tube 3, 1% acetic acid solution (post rinse swab). The staining procedure was carried out according to the manufacturer's instructions" Positivity threshold: positive result: stained dark or light blue; negative result: no stain Sequence of tests: VELscope then ViziLite then toluidine blue followed by the reference test Training or calibration of clinicians: not discussed Blinding of examiners: not discussed Multiple tests: yes; conducted independently during the same session, by the same examiners. Order: VELscope, ViziLite, toluidine blue Method of site selection: visual examination by 2 examiners Conflict of interests: equipment provided by manufacturers and marketers of toluidine blue kits, Zila Inc.
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "A surgical biopsy was performed for histopathological assessment... the presence or absence of dysplasia was further reviewed by an experienced oral pathologist"

Awan 2012 *(Continued)*

	Positivity threshold: quote: "The presence or absence of dysplasia in the biopsy specimen was recorded by two experienced oral pathologists" Sequence of tests: index tests followed by reference standard Training or calibration of pathologists: 2 experienced pathologists Blinding of examiners: unclear Multiple tests: no Method of site selection: the index test results were used to inform the selection of biopsy site Target condition: dysplasia
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: the tests were conducted during the same session for patient convenience Patients receiving both index and reference test but excluded from analysis: 24 patients receive toluidine blue test and reference test but were excluded from analysis, because of the characteristics of their lesions - patients with mixed red and white lesions were excluded from the analysis
Comparative	
User defined 1	
Notes	

Methodological quality

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

Awan 2012 *(Continued)*

knowledge of the results of the index tests?			
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Baeten 2018

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Subjects were recruited and screened at walk-in clinics," no details on methods of recruitment. The study separately included people with suspicious oral lesions (> 14 days), with no prior pathology and a second group with clinically normal oral mucosa (not considered further in this review). Suspicious clinical lesions included leukoplakia, erythroplakia, speckled leukoplakia, ulcerated lesions or masses present for > 14 days Exclusions: 5 participants due to methodological differences, 1 inconclusive biopsy, 6 did not consent to biopsy, 1 had a diagnosis which did not require biopsy, and 1 had a recent biopsy result
Patient characteristics and setting	Age: over 18 years Sex: 51 male, 18 female SES: not reported Ethnicity: not reported Stated risk factors: 20 tobacco smokers, 15 tobacco chewers, 4 both Number of patients/lesions: 69 participants enrolled, 55 included - 64 lesions (53 suspicious lesions and 11 normal) Lesion site: oral cavity Severity: suspicious lesions

Baeten 2018 *(Continued)*

	Country: India Type of facility: walk-in centre Prevalence: 36/53 lesions
Index tests	Category: vital staining - Wheat Germ Agglutinin-fluorescein isothiocyanate Description: quotes: "before WGA-FITC application, tissue autofluorescence was documented and photographed utilizing the excitation light source"... "The oral cavity was then cleansed via an oral rinse", followed by a distilled water rinse, "1 mL of 20 lM WGA-FITC solution was applied with a swab for approximately 10 seconds to observable oral lesions, including a 2-cm margin".. "subject then rinsed with approximately 30 mL of distilled water to remove any residual solution. After rinsing, the clinician used the excitation light 173 source and custom filter glasses to scan the oral cavity" Positivity threshold: level of fluorescence was graded on a scale of 0 to 4 (0 = no WGA-FITC fluorescence; 4 = very bright WGA-FITC fluorescence). Quote: "any aberrant staining between the suspicious lesion(s) and surrounding normal tissue was considered a 'positive' test result" Sequence of tests: index test followed by reference standard Training or calibration of clinicians: 2 study personnel graded the autofluorescence but no training or calibration was reported Blinding of examiners: not reported Multiple tests: no Method of site selection: observable oral lesions Conflict of interests: 2 employees of Visual Solutions
Target condition and reference standard(s)	Category: incisional or punch biopsy and histopathology Description: quote: "formalin-fixed, embedded in paraffin, microtome sectioned, and stained with hematoxylin - 205 eosin for histopathological evaluation" Positivity threshold: cancer, dysplasia, benign Sequence of tests: followed index test Training or calibration of pathologists: quote: "board certified examiner" Blinding of examiners: quote: "All specimens were analyzed by a blinded oral pathologist" Multiple tests: no Method of site selection: oral lesion Target condition: cancer and dysplasia
Flow and timing	Patients receiving index test but not reference test: 8 (6 did not consent to biopsy, 1 had a diagnosis which did not require biopsy, and 1 had a recent biopsy result) Patients receiving reference test but not index test: 0 Time interval: not reported, no reason to expect a time interval between tests Patients receiving both index and reference test but excluded from analysis: 1 - inconclusive
Comparative	
User defined 1	

Baeten 2018 *(Continued)*

Notes	We contacted study authors, they will confirm numbers for suspicious lesions only, without the normal mucosa which are included in the results of the paper

Methodological quality

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

Baeten 2018 *(Continued)*

Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	Low risk

Cancela-Rodriguez 2011

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "One hundred and sixty patients referred to the Department of Oral Medicine, Faculty of Dentistry, Complutense University of Madrid, Spain were selected for this study" "patients presented with 160 mucosal lesions, which required biopsy evaluation"
Patient characteristics and setting	Age: 55.3 +- 16.1 years, range 13 to 100 years Sex: 77 male, 83 female SES: not reported Ethnicity: all Caucasian Stated risk factors: 34% smokers, 27% consumed alcohol regularly Number of patients/legions: 160/160 Lesion site: oral cavity Severity: quote: "subjects with benign lesions or clinically suspicious premalignant or malignant lesions that were either white or red, exophytic or presenting as non-healing ulcers' Country: Spain Type of facility: secondary Prevalence: 29/160
Index tests	Category: vital staining - toluidine blue Description: quote: "Toluidine Blue was applied as a mouth rinse using the protocol described by Mashberg, with 1% aqueous acetic acid applied initially as a mucolytic agent and after Toluidine Blue rinsing to remove excess stain" Positivity threshold: quote: "The stain was considered positive when the surface mucosa took on a blue colour, either if the entire lesion was stained or just a portion of it. Those that do not took colouration or with equivocal findings were considered negatives" Sequence of tests: index followed by reference Training or calibration of clinicians: quotes: "a working clinical diagnosis was established using WHO Criteria 1980" "The test outcome was subjected to clinical evaluation, by four experienced oral pathologists previously calibrated in pairs" Blinding of examiners: index completed prior to reference Multiple tests: no Method of site selection: toluidine blue rinse so no site selection Conflict of interests: not reported

Cancela-Rodriguez 2011 *(Continued)*

Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: biopsy method unclear Positivity threshold: quote: "Following histopathological diagnosis, all lesions were classified in two groups: non-dysplastic/nonmalignant lesions, when there were no signs of dysplasia or histological malignancy and dysplastic/malignant lesions, when dysplasia or invasion was present" Sequence of tests: index followed by reference Training or calibration of pathologists: quote: "To avoid any inter examiner variability, the biopsies from this study were evaluated by the same pathologist to determine the presence and degree of dysplasia, or malignancy" Blinding of examiners: unclear whether the same pathologist evaluates the index and reference test Multiple tests: no Method of site selection: quote: "For lesions with a positive toluidine test, the biopsy was taken from the stained area" Target condition: dysplasia or histological malignancy
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: quote: "biopsy was taken from the stained area" so during the same appointment Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Cancela-Rodriguez 2011 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Chainani-Wu 2015

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "Patients presenting to the tertiary oral medicine referral clinic at UCSF for initial or follow-up evaluations, who were diagnosed with oral leukoplakia, erythroleukoplakia, or erythroplakia, were enrolled" "A consecutive sample of eligible, consenting participants was enrolled between November 2006 and March 2008"
Patient characteristics and setting	Age: 61 +- 10.6 years, range 42 to 90 years Sex: 23 male, 20 female SES: not reported Ethnicity: all Caucasian Stated risk factors: 64% current or past smokers, alcohol consumption not reported Number of patients/legions: 43/77 Lesion site: oral cavity

Chainani-Wu 2015 *(Continued)*

	Severity: quote: "patients with oral leuko(erythro)plakia, lesions clinically identified as 'premalignant' or 'suspicious for carcinoma'" Country: USA Type of facility: tertiary Prevalence: 70/77
Index tests	Category: light-based - ViziLite Description: quote: "patients were asked to rinse with a 1% acetic acid mouthrinse. In a darkened room, a chemiluminescent light was then used for visual inspection of the oral cavity" Positivity threshold: quote: "ViziLite examination that demonstrated increased brightness in comparison with the visual examination was described as ViziLite positive" Sequence of tests: index followed by reference Training or calibration of clinicians: quote: "the results recorded were based on agreement between both of these investigators" Blinding of examiners: index completed prior to reference Multiple tests: yes Method of site selection: not reported Conflict of interests: quote: "This study was funded by Zila Corporation" Category: light-based (ViziLite) plus staining (toluidine blue) Description: quote: "A preswab containing 1% acetic acid was applied over any visible lesions, and patients were then asked to rinse and expectorate using tap water. This was followed by application of a swab with 1% toluidine blue (tolonium chloride) on the lesions and then a postswab containing 1% acetic acid, and patients were again asked to rinse and expectorate using tap water" Positivity threshold: quotes: "The toluidine blue uptake in the lesions was recorded as positive or negative" "Any staining including a dark or pale blue mucosal uptake, including uptake in fissures, was classified as a positive response." For this combined test, a positive test result was indicated when both the ViziLite and the toluidine blue tests were positive Sequence of tests: index followed by reference Training or calibration of clinicians: quote: "the results recorded were based on agreement between both of these investigators" Blinding of examiners: index completed prior to reference Multiple tests: yes Method of site selection: quote: "application of a swab with 1% toluidine blue (tolonium chloride) on the lesions" Conflict of interests: quote: "This study was funded by Zila Corporation"
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quotes: "An incisional biopsy was carried out under local anesthesia, using either a punch or a scalpel depending on the site and extent of the lesion" "All biopsies were carried out in areas of clinically apparent white, red, or red-white change (these changes were noted on visual examination before toluidine blue application)"

Chainani-Wu 2015 *(Continued)*

	Positivity threshold: quote: "Following histopathological diagnosis, all lesions were classified in two groups: non-dysplastic/non-malignant lesions, when there were no signs of dysplasia or histological malignancy and dysplastic/malignant lesions, when dysplasia or invasion was present" Sequence of tests: index followed by reference Training or calibration of pathologists: quote "To avoid any inter examiner variability, the biopsies from this study were evaluated by the same pathologist to determine the presence and degree of dysplasia, or malignancy" Blinding of examiners: quote: "Information regarding toluidine blue staining was not provided to the pathologist" Multiple tests: no Method of site selection: quotes: "All biopsies were carried out in areas of clinically apparent white, red, or red-white change (these changes were noted on visual examination before toluidine blue application)" "the site of the incisional biopsies was determined based on identification of the most suspicious site within a lesion(s) based on clinical appearance and/or toluidine blue uptake by agreement between these two investigators" Target condition: dysplasia or histological malignancy
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: quote: "These procedures were all performed at a single clinic visit and were completed in approximately 1–2 h" Patients receiving both index and reference test but excluded from analysis: 7 benign lesions (3 non-specific ulcers, 1 lupus, 3 lichen planus) were excluded from the calculation of sensitivity and specificity
Comparative	
User defined 1	
Notes	Study used 3 tests, ViziLite, toluidine blue, and combined ViziLite and toluidine blue. This meta-analysis included results for ViziLite alone, or ViziLite and toluidine blue in combination. The results of toluidine blue alone were not included, as this test was not independent of the ViziLite test

Methodological quality

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

Chainani-Wu 2015 *(Continued)*

Are there concerns that the included patients and setting do not match 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?	No		
Could the patient flow have introduced bias?		High risk	

Chaudhari 2013

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "Phase II: Of the study all the cases detected positive by chief investigator and expert examiner were considered" "82 participants consented for staining procedure and biopsy procedure"
Patient characteristics and setting	Age: 34.98 years (SD 12.65), range 19 to 69 years Sex: all male SES: not reported Ethnicity: all Caucasian Stated risk factors: 353 (13.7%) smoked mostly Bidis, 702 (27.3%) were tobacco chewers, 1415 (55.01%) were both smokers as well as tobacco chewers; alcohol consumption not reported Number of patients/legions: 43/77 Lesion site: oral cavity Severity: all dysplasia classed as positive Country: India Type of facility: male prison Prevalence: 76/82 Note: these demographic details refer to the screening study (sample size 2572) from which the participants for the diagnostic study were drawn
Index tests	Category: staining (toluidine blue, Lugol's iodine) Description: quote: "1% toluidine blue and 2% Lugol's iodine were applied as illustrated by Epstein JB10" Positivity threshold: not explicitly stated, reported as + or - Sequence of tests: index followed by reference Training or calibration of clinicians: quote: "Training and calibration of the investigator (AC) in examination process was carried in the Department of Oral medicine and Radiology of the institution" Blinding of examiners: independent assessments of index tests. Quote: "Out of 82 participants, 41 participants were first screened with toluidine blue using above procedure and interpretation. After application of toluidine blue the 41 participants underwent Lugol's iodine application after 24 hours of washout period. All the procedure was interchanged for 41 participants i.e. 41 participants were first screened with Lugol's iodine and subsequently with toluidine blue. Lugol's Iodine was applied on the same region 24 hrs after toluidine blue application and subjected for biopsy" Multiple tests: yes Method of site selection: based on clinical examination Conflict of interests: no conflict of interest statement
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "After taking biopsy, specimen was immediately stored in 10% Formalin solution and subjected to histopathological analysis" Positivity threshold: quote: "For study purpose histological interpretation was limited to the dysplasia present (positive) or absent (negative)"

Chaudhari 2013 *(Continued)*

	Sequence of tests: index followed by reference Training or calibration of pathologists: quote: "All histopathological interpretations were carried out by the Oral Pathologist (SN)." Training and calibration not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: not reported Target condition: dysplasia or histological malignancy
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: quote: "Toluidine blue application and subjected to biopsy" Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			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		

Chaudhari 2013 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
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	

Chen 2007

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Fifty-eight patients who presented with suspected oral lesions were recruited into the study"
Patient characteristics and setting	Age: not reported Sex: not reported SES: not reported Ethnicity: not reported Stated risk factors: quote: "About two-thirds of cases consisted of smokers or people who chewed betel nuts" Number of patients/lesions: 58 patients and lesions Lesion site: oral cavity Severity: quote: "The lesions were characterised as homogeneous leukoplakia, heterogeneous leukoplakia, erythroplakia, and ulceration" Country: Taiwan Type of facility: secondary Prevalence: (cancer or precancerous) 29/58
Index tests	Category: vital staining - methylene blue

Chen 2007 *(Continued)*

	Description: quote: "A standard staining procedure included gargling with rinsing solution (1% lactic acid in purified water flavoured with raspberry) for 20 s, subsequently after power air spray on the lesion, the dye was given by gargling methylene blue (1% malachite, 1% eosin, 0.5% glycerol and dimethyl sulphoxide) for 20 s followed by an additional 20 s of gargling with the rinsing solution" Positivity threshold: unclear Sequence of tests: index followed by reference Training or calibration of clinicians: not reported. Quote: "The pattern and intensity of the staining was recorded and supervised by an experienced oral surgeon" Blinding of examiners: index completed prior to reference Multiple tests: no Method of site selection: methylene rinse so no site selection Conflict of interests: quote: "grant supported by NSC-94-2314B075 and VGH94242C"
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: biopsy method unclear Positivity threshold: quote: "Following histopathological diagnosis, all lesions were classified in two groups: non-dysplastic/non-malignant lesions, when there were no signs of dysplasia or histological malignancy and dysplastic/malignant lesions, when dysplasia or invasion was present" Sequence of tests: index followed by reference Training or calibration of pathologists: quote: "The pathological diagnosis was examined and verified independently by two specialists." Unclear Blinding of examiners: examinations were independent Multiple tests: no Method of site selection: quote: "The most obviously stained area or the most suspicious looking area, if there was no update of dye, was biopsied" Target condition: dysplasia and squamous cell carcinomas
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: completed at the same time Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Chen 2007 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 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	

Cheng 2003a

Study characteristics	
Patient Sampling	Method of patient selection: any patients with mucosa lesion referred to College of Stomatology, Sun Yat-sen University between October 2000 and March 2001 were included

Cheng 2003a *(Continued)*

Patient characteristics and setting	Age: 58.3 years (range 7 to 76) Sex: male 11%, female 89% SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/lesions: 60/60 Lesion site: mainly on buccal and lingual areas Severity: not stated Country: China Type of facility: College of Stomatology, Sun Yat-sen University Prevalence: 54/60
Index tests	Category: vital staining - Oratest Description: Oratest staining - similar to toluidine blue. Use the pre-exam mouthwash fluid for 20 seconds, wash the mouth twice with water, then use 1% toluidine blue to wash the mouth for 20 seconds and then wash the mouth with water twice, then use another mouthwash fluid (not clearly described) for 30 seconds. For the topical application group, 1% toluidine blue was used topically without mouth washing Positivity threshold: blue staining of the lesion predict a positive outcome, blurred blue staining which could not be washed out by the mouthwash fluid was also considered as positive Sequence of tests: before pathological test Training or calibration of clinicians: no training reported Blinding of examiners: index test completed before reference standard Multiple tests: no Method of site selection: not reported Conflict of interests: not reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: scalpel biopsy-based pathology, no further description provided Positivity threshold: not reported Sequence of tests: index then reference test Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: not reported Target condition: oral cancers and epithelial dysplasia
Flow and timing	Patients receiving index test but not reference test: 0

Cheng 2003a *(Continued)*

	Patients receiving reference test but not index test: 0 Time interval: not clear Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
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		

Cheng 2003a *(Continued)*

Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Unclear risk

Cheng 2003b

Study characteristics	
Patient Sampling	Second entry created for Cheng 2003, the 60 patients were split into 2 groups, one used a topical application and the other a rinse. To allow for covariate analysis a second entry is necessary to split the data for the results
Patient characteristics and setting	
Index tests	
Target condition and reference standard(s)	
Flow and timing	
Comparative	
User defined 1	
Notes	

Chiang 2019

Study characteristics	
Patient Sampling	Method of patient selection: quote: "150 consecutive patients with mucosal disorders," 10 excluded as they refused to consent to biopsy, 8 were unable due to racial protection law, 6 had lesions in oropharynx
Patient characteristics and setting	Age: 55.9 years (SD 12.7) Sex: 110 male, 16 female SES: not reported Ethnicity: not reported Stated risk factors: none stated Number of patients/lesions: 126 participants Lesion site: oral cavity, 71 buccal mucosa Severity: suspected lesions Country: Taiwan Type of facility: secondary care

Chiang 2019 *(Continued)*

	Prevalence: 63 dysplasia, 6 cancer, total 69/126
Index tests	Category: light-based - Hours UOC 100™ digital autofluorescence camera Description: quotes: "handheld 1920 x 1080-pixel camera designed to display mucosal changes on a 3.5-in. full-color TFT–LCD screen" "emits light in the 400–460-nm spectrum" Positivity threshold: quote: "findings were listed as fluorescence visualization loss (FVL) or fluorescence visualization retained (FVR)" Sequence of tests: index test followed by reference standard Training or calibration of clinicians: quote: "course taught by an experienced professional" Blinding of examiners: not reported, but index test performed first so assumed Multiple tests: no Method of site selection: oral lesions following COE carried out by a certified experienced examiner specialized in oral and maxillofacial surgery Conflict of interests: authors declare that they have no conflict of interest
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "A biopsy was performed for histopathological assessment" Positivity threshold: quote: "presence or absence of dysplasia or oral cancer" Sequence of tests: followed index test Training or calibration of pathologists: quote: "reports were confirmed and graded by a certified pathologist" Blinding of examiners: assume different examiners Multiple tests: no Method of site selection: quote: "any area at the abnormal mucosa with autofluorescence loss or retained" Target condition: dysplasia (oral cancer cases were excluded)
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: completed at the same time Patients receiving both index and reference test but excluded from analysis: 6 patients with oral cancer were excluded from the final analysis as the remit of the paper was to detect/diagnose OPMDs/dysplasia
Comparative	
User defined 1	
Notes	Numbers taken from Table 3 dysplasia rows, but does not include the 6 cancer diagnoses. We contacted authors for this information on 11th February 2020

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Chiang 2019 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match 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?	No		
Could the patient flow have introduced bias?		Low risk	

Delavarian 2010

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "study group consisted of 25 patients with 26 lesions which had been visited from Oct 2005 to Jan 2007, at Oral Medicine of Mashhad Faculty of Dentistry and Otorhinolaryngol-

Delavarian 2010 *(Continued)*

	ogy Department of QAEM, IMAM, REZA and OMID hospitals, Mishhad, Iran" "Inclusion criterion was: lesions clinically diagnosed as oral potentially malignant (leukoplakia, OLP) or malignant lesions (OSCC and verrucous carcinoma) and requiring an incisional biopsy for definite diagnosis"
Patient characteristics and setting	Age: 54.00 ± 17.38 years Sex: 13 male, 12 female SES: not reported Ethnicity: not reported Stated risk factors: unclear Number of patients/lesions: 25/26 Lesion site: oral cavity Severity: quote: "lesions clinically diagnosed as oral potentially malignant (leukoplakia, OLP) or malignant lesions (OSCC and verrucous carcinoma) and requiring incisional biopsy for definite diagnosis" Country: Iran Type of facility: secondary Prevalence: (dysplasia/malignancy) 8/26 Exclusions: history of any treatment for the lesion, systemic contraindication for scalpel biopsy
Index tests	Category: cytology - OralCDx (laboratory processing not performed at OralCDx laboratory) Description: quote: "After determination of site biopsy, under local anaesthesia, needed for scalpel biopsy, the OralCDx brush was placed in the selected area and turned 5 to 10 times until appearing pinpoint bleeding-upon manufacture's recommendation" Positivity threshold: quote: "The pathological findings were categorized as 3 groups: 1) Positive: dysplastic epithelial changes, 2) Negative: absence of any evidence suggesting dysplasia, 3) Inadequate sampling" Sequence of tests: index followed by reference Training or calibration of clinicians: unclear, quote: "They were examined by a pathologist informed about clinical diagnosis" Blinding of examiners: index completed prior to reference, quote: "blind to the histopathological results" Multiple tests: no Method of site selection: quote: "The most impressive site of biopsy was determined upon one if these criteria: 1) The most probable site of dysplasia/malignancy OR 2) High risk area for dysplasia/malignancy OR 3) The most surgically accessible site" Conflict of interests: university support acknowledged
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: scalpel biopsy method unclear Positivity threshold: quote: "The Pindborg criteria for detecting dysplasia and malignancy were used and the histopathologic diagnosis was made. The presence of dysplasia/malignancy in histopathology was classified as normal, mild, moderate and severe dysplasia (level 1 to 3), carcinoma in situ (level 4) and carcinoma (level 5)" Sequence of tests: index followed by reference

Delavarian 2010 *(Continued)*

	Training or calibration of pathologists: unclear Blinding of examiners: quote: "The histopathologic preparations were observed by the same pathologist blind to the cytopathical study and informed about clinical diagnosis" Multiple tests: not applicable Method of site selection: quote: "The scalpel biopsy was done immediately in the site of pin-point bleeding" Target condition: quote: "The Pindborg criteria for detecting dysplasia and malignancy were used and the histopathologic diagnosis was made. The presence of dysplasia/malignancy in histopathology was classified as normal, mild, moderate and severe dysplasia (level 1 to 3), carcinoma in situ (level 4) and carcinoma (level 5)"
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not discussed Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Assumed figures in Table 2 should be reversed: 'Gold Standard' Normal and Disease should be reversed. This has been done for data entry

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Delavarian 2010 *(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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	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	

Du 2007

Study characteristics	
Patient Sampling	Method of patient selection: 132 patients from School and Hospital of Stomatology, Wuhan University, China between July 2002 and November 2003 Inclusion criteria for patients: (i) superficial ulceration suspicious of malignancy (ii) oral leukoplakia (iii) oral lichen planus (iv) oral leukokeratosis Excluded: (i) benign oral lesion (ii) lesion without histological result after clinical diagnosis

Du 2007 *(Continued)*

Patient characteristics and setting	Age: mean 49.4 (SD 12.7) years Sex: 67 female, 61 male SES: not specified Ethnicity: not specified Stated risk factors: not discussed Number of patients/lesions: 128/128 Lesion site: oral cavity Severity: all patients suspected of malignancy, leukoplakia, lichen planus, leukokeratosis Country: China Type of facility: secondary Prevalence: 33/128
Index tests	Category: vital staining - rose bengal Description: quote: "(i) oral examination of the location, size, morphology and surface characteristics of the lesions, (ii) mouth rinse with distilled water (reagent A) to clean the lesions for 1 minute, (iii) applications of solution of RB (reagent B) with cotton tip for 2 minutes, (iv) mouth rinse with distilled water (reagent C) to remove excess RB solution for 1 minutes, (v) oral examination of the location, size, morphology and surface characteristics of sites stained" Positivity threshold: quote: "Staining result of a lesion was classified as 1, 2, 3 or 4 according to the shade tabs. In the present study, staining results of 3 and 4 were regarded as RB positive staining, while staining results of 1 and 2 were regarded as RB negative staining" Sequence of tests: unclear regarding ordering, implied that histology performed after index Training or calibration of clinicians: single clinician training not reported Blinding of examiners: assume given implied sequence of tests Multiple tests: no Method of site selection: oral examination Conflict of interests: grant sponsor, Science and Technology Bureau of Wuhan City, People's Republic of China; grant number: 20026002084, assume no conflict
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: scalpel used, methods unclear Positivity threshold: unclear reporting, but outcomes reported as either malignant (including dysplasia or squamous cell carcinoma) or benign Sequence of tests: unclear, but assumed after index from reporting structure Training or calibration of pathologists: none, only 1 experienced oral pathologist used Blinding of examiners: blinded to index test Multiple tests: no Method of site selection: unclear but each patient had 1 lesion, assume biopsy taken from this lesion

Du 2007 *(Continued)*

	Target condition: malignant or benign
Flow and timing	Patients receiving index test but not reference test: 4 Patients receiving reference test but not index test: 0 Time interval: not specified Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Du 2007 *(Continued)*

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

Epstein 2008

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "Patients were identified in clinics at three study sites: the University of California San Francisco (UCSF), the University of Illinois Chicago (UIC), and the British Columbia Cancer Agency (BCCA)" "Patients who had a history of oral lesions or were at high risk for an oral lesion were identified and asked to participate"
Patient characteristics and setting	Age: mean 59.4 (SD 12.53) years Sex: male 48.81%, female 51.19% SES: not discussed Ethnicity: not discussed Stated risk factors: cigarette smokers: current 15.48%, prior 41.67% alcohol consumers: 2/3 cigar/pipe/chew/snuff: less than 10% no betel nut Number of patients/lesions: 84/97 Lesion site: not reported, but was recorded by examiners Severity: patients identified with a lesion on conventional visual examination Country: USA Type of facility: secondary care Prevalence: 42/97
Index tests	Category: vital staining plus adjunct - ViziLite and toluidine blue Description: lesions were identified by visual examination, acetic acid rinse applied for 30 to 60 seconds, then evaluated by chemiluminescent light, lesion then swabbed with acetic acid solution followed by swabbing of toluidine blue, then a further swab of acetic acid Positivity threshold: quote: "The investigator reported their subjective assessment of the impact of chemiluminescence upon lesions characteristics of brightness, sharpness, surface texture, and/or size

Epstein 2008 *(Continued)*

	using a four point Likert scale (decreased, no change, slight improvement, marked improvement). After the toluidine blue staining the investigator recorded the staining pattern either as negative, incomplete, or complete total lesion staining." Potential confusion over "incomplete" Sequence of tests: visual, light-based, vital stain then reference test Training or calibration of clinicians: not discussed Blinding of examiners: not blinded of other index tests, index test defined as the combination of light-based and vital stain Multiple tests: yes. Combined test results used (Table 4 toluidine blue) as authors report 'adjunct' tests and cumulative values Method of site selection: from oral examination Conflict of interests: funded by Trylon Corp, authors linked to Zila Inc.
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "All visualized lesions were then biopsied using either scalpel or punch technique. All procedures were conducted within a single patient visit. Histopathologic diagnoses were completed by board certified Oral Pathologists who were blinded to the clinical findings" Positivity threshold: quote: "Serious pathology was selected to refer to severe dysplasia, CIS and frank SCC. Benign (referring to microscopic evidence of no epithelial dysplasia), mild and moderate dysplasias were classified as non-serious pathology for this analysis" Sequence of tests: index then reference Training or calibration of pathologists: not discussed, although quote "board qualified oral pathologist" Blinding of examiners: yes Multiple tests: no Method of site selection: all lesions visualised during index test Target condition: severe dysplasia, CIS and SCC
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: minimal, completed on the same day Patients receiving both index and reference test but excluded from analysis: 1 will not have significant effect on results
Comparative	
User defined 1	
Notes	Concerns regarding this study analysis for cancer only, ignoring OPMDs. Positive results include only severe dysplasia. We have re-classified the reported data with a positivity threshold of mild dysplasia and above. Results screened equivocally have been categorised as positive

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Epstein 2008 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Farah 2007

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Fifty-five patients referred to an oral medicine specialist service over a 3 month period for assessment of an oral mucosal white lesion were prospectively screened with ViziLite"
Patient characteristics and setting	Setting: oral medicine specialist, secondary care Age: male 56.81 (2.2), female 58.7 (2.47) years Sex: male 47%, female 53% SES: not discussed Ethnicity: not discussed Stated risk factors: 21 smokers (38%) Number of patients/lesions: 55/80, only reported on primary lesions 55 Lesion site: alveolar ridge: primary 10/satellite 2; tongue: 17/7; buccal mucosa: 17/13; floor of mouth: 5/1; gingiva: 2/1; palate: 2/1; lip: 2/0 Severity: oral mucosal white lesion Country: Australia Type of facility: secondary Prevalence: 10/55
Index tests	Category: light-based - ViziLite Description: quote: "After a 60 s rinse with 1% acetic acid solution, the outer flexible capsule of the light stick was bent, breaking the inner fragile glass vial. The light stick was shaken vigorously to mix the contents, and then placed in the open end of the retractor and assembled. The room and operatory lights were dimmed, and the examination with ViziLite illumination was undertaken" Positivity threshold: unclear, although the authors do state "all lesions appeared 'aceto-white' under chemiluminescent light, and were considered ViziLite positive." It is not clear that this detail was used in the diagnostic decision Sequence of tests: index then reference Training or calibration of clinicians: quote: "calibration amongst the two oral medicine specialists." Unclear whether this refers to the decision on the lesion or prior training/calibration Blinding of examiners: unclear Multiple tests: no Method of site selection: intraoral examinations Conflict of interests: none
Target condition and reference standard(s)	Category: biopsy with histopathological assessment Description: quote: "an incisional scalpel biopsy was performed under local anaesthesia to obtain a definitive histopathological diagnosis"

Farah 2007 *(Continued)*

	Positivity threshold: taken from table 3, results of histology worked back to clinical definitions: homogenous leukoplakia/keratosis, fibroepithelial hyperplasia, oral lichen planus/lichenoid stomatitis, non-homogenous leukoplakia/epithelial dysplasia, squamous cell carcinoma, non-specific ulceration Sequence of tests: index then reference Training or calibration of pathologists: not discussed Blinding of examiners: quote: "registered oral pathologists not involved with the clinical study" Multiple tests: no Method of site selection: not discussed Target condition: see above
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: scalpel biopsy performed after clinical diagnosis Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Quote: "examination of the oral tissues with ViziLite illumination did not change the provisional diagnosis, nor alter the biopsy site"

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Farah 2007 *(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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Farah 2012

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "Patients presenting to an oral medicine specialist unit for assessment of an oral mucosal lesion were recruited into the study" "Patients known to have oral epithelial dysplasia or squamous cell carcinoma were not included in this study"
Patient characteristics and setting	Setting: oral medicine specialist, secondary care Age: mean - male 57.8 (+/- 11.88); female 59.08 (+/- 12.8) years Sex: male 46 (41.1%), female 66 (58.9%) SES: not reported Ethnicity: not reported Stated risk factors: smoker: 54/112, 48.2% alcohol consumer: 47/112, 42% smoker and alcohol consumer: 31/112, 27.7% Number of patients/lesions: 112/118

Farah 2012 *(Continued)*

	Lesion site: oral cavity Severity: quote: "oral mucosal white or mixed red/white lesion that was deemed [...] to be clinically suspicious" Country: Australia Type of facility: quote: "oral medicine specialist unit" Prevalence: 27/118 (3/118 OSCC; 24/118 PMD)
Index tests	Category: light-based - VELscope Description: quote: "Clinical examination was repeated using VELscope while the room and operatory lights were dimmed, and all measurements repeated" Positivity threshold: quote: "Lesions that showed loss of autofluorescence were deemed positive, and lesions that did not show any loss of autofluorescence were deemed negative. In addition, all lesions that lost autofluorescence were blanched to evaluate diascopic fluorescence, and those that were deemed negative for loss of autofluorescence only if complete blanching was achieved" Sequence of tests: index test followed by reference standard Training or calibration of clinicians: calibration of VELscope findings are reported, but no reporting of training. Quotes: "Clinical interobserver agreement was calculated with 71.4% agreement on the clinical provisional diagnosis, 60.7% agreement on the VELscope provisional diagnosis, 82% agreement on loss of autofluorescence, and 62% agreement on complete blanching after loss of autofluorescence" "Calibration of clinical observations was also undertaken between 2 of the authors (C.S.F. and M.J.M.) on a separate cohort of patients not included in this study" Blinding of examiners: index test completed prior to reference standard. Nothing specifically reported Multiple tests: no Method of site selection: visual examination Conflict of interests: not specifically reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessments Description: quotes: "A scalpel tissue biopsy was taken" "Biopsy specimens were fixed in formalin, blocked in paraffin, stained with hematoxylineosin, and assessed by routine histopathology" Positivity threshold: taken from table 4, results of histology worked back to clinical definitions Sequence of tests: index test followed by reference standard Training or calibration of pathologists: quote: "There was a 96.6% interobserver agreement for histopathological interpretation between the 2 pathologists" Blinding of examiners: quote: "Both examiners were blinded to the clinical findings" Multiple tests: no Method of site selection: quote: "Loss of autofluorescence was used to determine the best site for biopsy" Target condition: dysplasia and OSCC
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported

Farah 2012 *(Continued)*

Patients receiving both index and reference test but excluded from analysis: 0

Comparative			
User defined 1			
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		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 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?	Unclear		

Farah 2012 *(Continued)*

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

Fontes 2013

Study characteristics	
Patient Sampling	Method of patient selection: quote: "The study sample consisted of 172 patients with oral lesions clinically suspicious of malignancy who sought treatment in the outpatient clinic of oral diagnosis at Antônio Pedro University Hospital/UFF, Niterói, Brazil, and from other outpatient clinics enrolled in this study from 2002 to 2010"
Patient characteristics and setting	Age: range 20 to 93 years Sex: 114 male, 58 female SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/legions: 172/172 Lesion site: oral cavity Severity: quote: "oral lesions clinically suspicious of malignancy" Country: Brazil Type of facility: secondary Prevalence: 157/164
Index tests	Category: cytology - cytobrush device Description: quote: "the oral lesions were scraped with a cytobrush device by applying pressure and rotation. The cells were immediately smeared on a clean frosted glass slide and fixed in 95% ethanol" Positivity threshold: quote: "Biopsy specimens were embedded in paraffin, and 5-µm thick sections were obtained from paraffin blocks and stained with hematoxylin-eosin (H&E) according to the protocol established by the anatomic pathology service of the Antônio Pedro University Hospital/UFF" Sequence of tests: index followed by reference Training or calibration of clinicians: not reported Blinding of examiners: index completed prior to reference Multiple tests: no Method of site selection: not reported

Fontes 2013 *(Continued)*

	Conflict of interests: quote: "This study was supported by grants from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), a Brazilian governmental institution"
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quotes: "an incisional biopsy was performed, and the specimen was fixed in 10% formalin.." "Biopsy specimens were embedded in paraffin, and 5-µm thick sections were obtained from paraffin blocks and stained with hematoxylin-eosin (H&E) according to the protocol established by the anatomic pathology service of the Antônio Pedro University Hospital/UFF" Positivity threshold: quote: "Slides containing histopathological sections were evaluated according to the morphologic criteria established by the World Health Organization" Sequence of tests: index followed by reference Training or calibration of pathologists: quote: "The smears were evaluated at different times by three independent pathologists, and discordant results were reviewed and discussed until a consensus was reached" Blinding of examiners: not reported Multiple tests: no Method of site selection: not reported Target condition: dysplasia or histological malignancy
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported but assumed to be at the same appointment Patients receiving both index and reference test but excluded from analysis: 8, quote: "inadequate material for cytology analysis"
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear

Fontes 2013 *(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	

Ganga 2017

Study characteristics	
Patient Sampling	Method of patient selection: patients with oral mucosal lesions, possibly consecutive but not clearly reported
Patient characteristics and setting	Age: not reported Sex: not reported SES: not reported Ethnicity: not reported Stated risk factors: none stated Number of patients/lesions: 200 patients/lesions Lesion site: oral cavity

Ganga 2017 *(Continued)*

	Severity: quote: "Patients with oral mucosal lesions" Country: India Type of facility: Department of Oral Pathology and Microbiology Prevalence: 25/200 were malignant, the remainder were benign
Index tests	Category: light-based - VELscope Description: not reported Positivity threshold: loss of autofluorescence, quotes: "appeared dark compared to the surrounding unaltered tissue", versus "lesions that exhibited retention of autofluorescence" "Only a complete FVL was rated as malignant or dysplastic ...and lesions that demonstrated autofluorescence patterns other than a complete FVL were included in the FVR group" Sequence of tests: index followed by reference standard Training or calibration of clinicians: not reported Blinding of examiners: not reported, although ordering ensured blinding to histology Multiple tests: no Method of site selection: unclear Conflict of interests: authors declared no conflict of interest
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "Hematoxylin and eosin staining of the formalin fixed paraffin embedded tissue sections was carried out" Positivity threshold: benign/malignant - no dysplasia labelled in results Sequence of tests: index test followed by reference standard Training or calibration of pathologists: quote: "two experienced oral pathologists" Blinding of examiners: quote: "blinded to the VELscope findings" Multiple tests: no Method of site selection: oral lesions detected on COE Target condition: malignant/dysplastic lesions
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported, no reason to expect a time interval between tests Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Uncertain over other conditions which authors classify as benign but may be categorised as dysplasia in other studies, needs clarification (table 3)
Methodological quality	

Ganga 2017 *(Continued)*

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

Guneri 2011

Study characteristics

Guneri 2011 *(Continued)*

Patient Sampling	Method of patient selection: quote: "Thirty-five patients with oral mucosal lesions identified by the Orofacial Lesions Council of Ege University, Izmir, Turkey, were seen for further evaluation"
Patient characteristics and setting	Age: mean 56.2 years Sex: 13 male, 22 female SES: not stated Ethnicity: not stated Stated risk factors: 47% smokers Number of patients/lesions: 35/43 Lesion site: buccal mucosa 56%, tongue 19%, hard palate 14% Severity: quote: "Lesions selected for further examination with Tblue staining and brush cytology were homogenous and non-homogenous leukoplakia, reticular erosive/ulcerated lichenoid lesions, and superficial ulcerations suspicious of malignancy" Country: Turkey Type of facility: university clinic Prevalence: 15/43
Index tests	**Index test 1: toluidine blue** Category: vital staining Description: quote: "head, neck and intraoral examinations were completed before Tblue application, oral brush cytology and scalpel biopsy" Toluidine blue: quote: "The oral rinsing protocol was: 20 s pre-rinse with 30 ml of 1% acetic acid; 20 s water rinse; 20 s rinse/gargle with 10 ml of the 1% tolonium chloride solution; 20 s post-rinse with 30 ml of 1% acetic acid (twice); a final water rinse" Positivity threshold: quote: "The pattern of dye retention and the intensity of stain retention were recorded (2, dark blue staining; 1, minimal blue staining; 0, no blue staining). Occasionally, normal mucosa also appeared light blue, but this staining was not interpreted as positive" Sequence of tests: staining, brush biopsy, followed by reference standard Training or calibration of clinicians: not reported, although performed by 1 examiner Blinding of examiners: not specified Multiple tests: yes. Combined test results used (Table 1) as tests not carried out independently; results of first index test used to inform second index test Method of site selection: not reported Conflict of interests: stated no conflict and funded by university **Index test 2: brush cytology** Category: cytology - cytobrush Description: quote: "Brush cytology was performed using a Cytobrush Plus GT which was rotated on the lesion site with pressure, until pinpoint bleeding was observed. The harvested cells were transferred to a slide by a 3608 turning and rolling motion with the brush and the slides were washed rapidly with ethyl alcohol for fixation. Cytology specimens were stained with hematoxylin–eosin, and examined by an oral pathologist"

Guneri 2011 *(Continued)*

	Positivity threshold: quote: "The brush cytology results were classified as malignant, atypical (suspicious), benign tissues or inadequate sample" Sequence of tests: toluidine blue, brush biopsy, followed by reference standard Training or calibration of clinicians: not reported Blinding of examiners: not specified Multiple tests: yes. Combined test results used (Table 1) as tests not carried out independently; results of first index test used to inform second index test Method of site selection: not reported Conflict of interests: stated no conflict and funded by university
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quotes: "All lesions were subject to scalpel biopsy with selection based on the clinical appearance of the lesion" "The scalpel biopsy specimens were submitted in formalin for hematoxylin-eosin staining. After embedding, 5 mm thick sections were prepared" Positivity threshold: quotes: "Pathologic interpretation was based on established criteria and classified as squamous cell carcinoma, epithelial dysplasia, hyperkeratosis, lichen planus, and other benign" "biopsies confirmed as severe dysplasia, carcinoma-in-situ or SCC were considered 'positive'; no dysplasia, mild and moderate dysplasia were 'negative'" Sequence of tests: index then reference Training or calibration of pathologists: quote: "Surgical biopsies were performed by an experienced oral and maxillofacial surgeon" Blinding of examiners: not reported Multiple tests: 1 reference standard Method of site selection: quote: "All areas retaining Tblue were biopsied; in sites with no retention of staining, clinical judgment guided the biopsy procedure" Target condition: quote: "Pathologic interpretation was based on established criteria (WHO) and classified as squamous cell carcinoma, epithelial dysplasia, hyperkeratosis, lichen planus, and other benign lesions"
Flow and timing	Patients receiving index tests but not reference test: 0 Patients receiving reference test but not index tests: 0 Time interval: not more than 2 weeks between the 3 methods of investigation Patients receiving both index and reference test but excluded from analysis: 1 for brush biopsy considered a false positive because of inadequate cell sampling
Comparative	
User defined 1	
Notes	Toluidine blue prepared as an oral rinse since there is no pharmaceutical grade toluidine blue available in Turkey

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Guneri 2011 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Guneri 2011 *(Continued)*

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

Gupta 2007

Study characteristics	
Patient Sampling	Method of patient selection: quote: "96 patients with suspicious oral lesions who attended the outpatients clinics of Otorhinolaryngology Department, Swaroop Rani Nehru Hospital Allahabad, were screened"
Patient characteristics and setting	Age: (benign and premalignant) 38, range 19 to 75 years; (SCC) 52, range 35 to 74 years Sex: (benign and premalignant) 42 male, 22 female; (SCC) 24 male, 8 female SES: not reported Ethnicity: not reported (India) Stated risk factors: (benign and premalignant) 62% smokers, 66% chewed tobacco, 44% chewed pan, 47% consumed alcohol regularly; (SCC) 23% smokers, 80% chewed tobacco, 83% chewed pan, 19% consumed alcohol regularly Number of patients/lesions: 96 Lesion site: oral cavity Severity: quote: "suspicious premalignant or malignant lesions of the oral cavity irrespective of site, stage and sex were selected" Country: India Type of facility: secondary Prevalence: (premalignant and malignant) 46/96
Index tests	**Index test 1: toluidine blue** Category: vital staining Description: quote: "One percent aqueous toluidine blue was applied to suspicious lesion for 30 seconds. It was followed by a tap water or normal saline rinse. Then it was rinsed by 1% acetic acid for 30 seconds to reduce background staining" Positivity threshold: participants' results classed as positive or negative but no thresholds or inadequate/equivocal results reported. Sequence of tests: staining, brush biopsy, followed by reference standard Training or calibration of clinicians: not reported Blinding of examiners: iIndex completed prior to reference

Gupta 2007 *(Continued)*

	Multiple tests: yes. Combined test results used (Table 3) results of first index test used to inform second index test Method of site selection: not reported Conflict of interests: not reported **Index test 2: brush cytology** Category: cytology Description: quote: "Utilizing a small, hard toothbrush, a transepithelial biopsy was taken with minimum discomfort to the patient. To obtain an adequate specimen, moderate pressure was applied on the lesion by the brush until pinpoint bleeding was noted signalling entry into lamina propria and thus signalling entry into the full thickness of the epithelium" Positivity threshold: quote: "The following parameters were analysed in the smear: enlarged nuclei, variation in nuclear size and shape (pleomorphism), nuclear borders, nuclear/cytoplasmic ratio, number of nuclei, hyperchromatism, chromatin patter and distribution, and discrepancy in maturation." Participants' results classed as positive or negative but no thresholds or inadequate/equivocal results reported Sequence of tests: toluidine blue, brush biopsy, followed by reference standard Training or calibration of clinicians: not reported Blinding of examiners: index completed prior to reference. Not reported whether results of toluidine blue staining were known prior to biopsy Multiple tests: yes. Combined test results used (Table 3) results of first index test used to inform second index test Method of site selection: not reported Conflict of interests: not reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "For histology, after routine processing and paraffin embedding, several sections (3-4 mm in thickness) were cut from each case. At least 1 section of each case was stained with hematoxylin-eosin and examined for histological diagnosis" Positivity threshold: results reported as SCC, dysplasia or benign Sequence of tests: both index tests were followed by reference standard Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: not reported Target condition: benign (including oral submucous fibrosis and angiomatous malformation), premalignant (dysplasia) and squamous cell carcinoma
Flow and timing	Patients receiving index tests but not reference test: 0 Patients receiving reference test but not index tests: 0 Time interval: not reported Patients receiving both index and reference test but excluded from analysis: 0 All patients had toluidine blue staining, oral brush biopsy, and scalpel biopsy

Gupta 2007 *(Continued)*

Comparative	
User defined 1	
Notes	Excluded brush biopsy as an independent index test (blue staining present) and only used results for toluidine blue staining and combined index test of staining and brush biopsy. Numbers for true positive, false positive, true negative, and false negative have been calculated from raw data presented in table 3 sensitivity and specificity values. Results have been entered for both premalignant and malignant

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not			Unclear

Gupta 2007 *(Continued)*

match the question?		
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

Hanken 2013

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Patients presenting in the department of oral and maxillofacial surgery in the University Medical Center Hamburg Eppendorf to rule out invasive squamous cell carcinoma were recruited for this study in a prospective single blinded design"
Patient characteristics and setting	Age: 41 to 76 years Sex: 25 male, 35 female SES: not reported Ethnicity: not reported Stated risk factors: 88% current or previous smokers, 68% consumed > 20 g per day alcohol regularly Number of patients/lesions: 60/60 Lesion site: oral cavity Severity: dysplastic and malignant oral mucosa lesions Country: Germany Type of facility: secondary Prevalence: (premalignant and malignant) 48/60
Index tests	Category: light-based Description: quote: "The device was used under a dimmed room light ... The autofluorescence excitation device uses visible light in the 430 nm wave length" Positivity threshold: quote: "the complete loss of the normal tissue fluorescence (fluorescence visualization loss) was rated as malignant or dysplastic"

Hanken 2013 *(Continued)*

	Sequence of tests: index followed by reference standard Training or calibration of clinicians: quote: "Additionally both examiners were calibrated for the device in advance" Blinding of examiners: index completed prior to reference Multiple tests: no Method of site selection: not reported Conflict of interests: the authors declare that they have no competing interest
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "All specimens were placed in 4% buffered formalin solution for fixation. Paraffin embedded material was cut into 4 µm thick sections and stained with hematoxylin + eosin (H.E.; MERCK Germany). After color staining all mucosa lesions underwent a histo- pathological evaluation by two independent experienced pathologists" Positivity threshold: results reported as SCC, dysplasia or benign Sequence of tests: index test followed by reference standard Training or calibration of pathologists: quote: "two independent experienced pathologists" Blinding of examiners: quote: "two independent experienced pathologists" Multiple tests: no Method of site selection: not reported Target condition: dysplasia, carcinoma
Flow and timing	Patients receiving index tests but not reference test: 0 Patients receiving reference test but not index tests: 0 Time interval: not reported but assumed at the same appointment Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	

Methodological quality

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

Hanken 2013 *(Continued)*

Are there concerns that the included patients and setting do not match 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	

Jayasinghe 2020

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Patients who were above 18 years of age, with a clinical diagnosis of leukoplakia, erythroplakia, OSF, DLE or OLP (OPMDs) were recruited for the study"
Patient characteristics and setting	Age: not reported Sex: not reported SES: not reported Ethnicity: not reported Stated risk factors: not reported, only that patients with previous history of OSCC were excluded Number of patients/lesions: 60/60

Jayasinghe 2020 *(Continued)*

	Lesion site: oral cavity Severity: dysplastic and malignant oral mucosa lesions Country: Sri Lanka Type of facility: secondary Prevalence: (premalignant and malignant) 41/60
Index tests	Category: vital staining Description: quote: "1% TB is applied in 4 steps; first, rinsing of oral cavity with water for 20 s to remove debris, then application or rinse with 1% acetic acid for 20 s, after that application or rinse with TB solution (1% W/W) for 20 s and finally application or rinse with 1% acetic acid for 20 s to eliminate mechanically retained stains" Positivity threshold: quote: "A royal blue or a navy blue stain as shown in Fig. 1 was considered as a positive result even if either the entire lesion or a portion of the lesion is stained or stippled. A light blue staining is considered doubtful. If there is no colour absorbed by the lesion, it was taken as a negative stain" Sequence of tests: index followed by reference standard Training or calibration of clinicians: quote: "Dental surgeons in the selected institutions were trained by the expert in the field of Oral Medicine at their institutions" Blinding of examiners: index completed prior to reference Multiple tests: no Method of site selection: not reported Conflict of interests: not reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "surgical incisional biopsy was performed from all subjects for histopathological diagnosis and cases with positive TB test; the biopsy was performed from positive region under local anesthesia. Patients were managed with the standard management protocol (Fig. 2)" Positivity threshold: results reported as dysplasia or no dysplasia. Histopathological diagnosis was done according to WHO 2017 as mild moderate and severe epithelial dysplasia Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: quote: "the biopsy was performed from positive region" Target condition: dysplasia or no dysplasia
Flow and timing	Patients receiving index tests but not reference test: 5 Patients receiving reference test but not index tests: 0 Time interval: not reported but assumed at the same appointment Patients receiving both index and reference test but excluded from analysis: 0
Comparative	

Jayasinghe 2020 *(Continued)*

User defined 1			
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		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Jayasinghe 2020 *(Continued)*

Could the patient flow have introduced bias?	Low risk

Johnson 2019

Study characteristics	
Patient Sampling	Method of patient selection: not reported, reports recruitment at referral clinic and walk-in centres but no further details Exclusions: 2 had previous biopsy, 9 did not consent, 6 for "methodologic variances by study personnel", 3 had inconclusive pathology reports
Patient characteristics and setting	Age: over 18 years Sex: 62 male, 35 female SES: not reported Ethnicity: not reported Stated risk factors: tobacco users 27%, alcohol 9%, both 23% Number of patients/lesions: 117 patients recruited, 97 included - 86 had 100 clinically suspicious lesions Lesion site: oral cavity Severity: quote: "different suspicious oral lesions at various stages" Country: US and India Type of facility: quote: "a multisite clinical trial was performed at the North Memorial Hospital (Minneapolis, MN) and 2 walk-in clinics, the Mazumdar-Shaw Cancer Center and the KLES Dental College (Bangalore, India)" Prevalence: 66/100 were cancerous or dysplastic
Index tests	Category: combined system of light and stain - nuclear stain with UV-A fluorescence. Where positive results were observed wheat germ agglutinin conjugated to fluorescein isothiocyanate (WGA-FITC) was applied as a second test Description: quotes: "an ethanol based solution, with citric acid and surfactants, was used to clean the suspicious lesion(s) with a swab, including a 2-cm margin, followed by a water rinse" "a UV-excitable fluorescent nuclear stain was applied to the suspicious lesion(s) and margin using a swab, followed by another water rinse." A UVA light was used to assess staining Positivity threshold: quote: "If the nuclear stain's fluorescent pattern was uniform between the lesion(s) and the margin, then the system's result was 'negative'" Sequence of tests: clinical exam with UV-A autofluorescence, nuclear stating, WGA-FITC, then reference standard Training or calibration of clinicians: not reported Blinding of examiners: not reported Multiple tests: no (combined test as described above) Method of site selection: oral lesions

Johnson 2019 *(Continued)*

	Conflict of interests: 2 authors employed by Inter-Med, Inc, Racine, WI
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment (brush biopsies used for healthy control) Description: quotes: "scalpel or 5-mm punch biopsy" "fixed in formalin, embedded in paraffin, sectioned by microtome, and stained with hematoxylin and eosin (H&E) for diagnosis" Positivity threshold: cancerous, dysplastic, or benign Sequence of tests: index tests followed by reference standard Training or calibration of pathologists: quote: "board-certified oral pathologist" Blinding of examiners: yes Multiple tests: no Method of site selection: oral lesions Target condition: cancerous/dysplastic
Flow and timing	Patients receiving index test but not reference test: 9 Patients receiving reference test but not index test: 0 Time interval: not reported, no reason to expect a time interval between tests Patients receiving both index and reference test but excluded from analysis: 3 had inconclusive pathology reports
Comparative	
User defined 1	
Notes	Demographics reported for Groups 1 and 2 combined - not reported separately in the manuscript for Group 1 only. For the purposes of meta-analysis we have taken the results of Table 2 and removed the results for the 11 normal controls that all tested negative

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Johnson 2019 *(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?	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?		High risk	

Kämmerer 2013

Study characteristics	
Patient Sampling	Method of patient selection: quote: "From 01/2005 to 01/2009, 70 patients (45 male, 25 female; mean age 62 years (range 27–88)) with 88 oral lesions of uncertain dignity were included in the study"
Patient characteristics and setting	Age: mean age 62 years, range 27 to 88 years Sex: 45 male, 25 female SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/lesions: 70/88 Lesion site: oral cavity Severity: quote: "clinically suspicious but not obvious malignant lesions"

Kämmerer 2013 *(Continued)*

Country: Germany

Type of facility: secondary

Prevalence: (premalignant and malignant) 37/51

Index tests

Index test 1 - brush cytology

Category: cytology - cytobrush

Description: quote: "From each lesion, four slides from a separate brush biopsy were taken without prior mucosal disinfection. A proprietary brush (Cytobrush Plus GT; Medscan, Malmo, Sweden) was used to obtain epithelial cells. The small brush was rotated for 10 revolutions over an approximately 10 x 10 mm (100 mm2) area, resulting in most cases in petechial bleeding points. Immediately after harvesting, the cells were transferred onto four pre-charged glass microscope slide (SuperFrost Plus, Menzel, Braunschweig, Germany) in a rotating manner and fixed by ethyl alcohol (Merckofix, Merck, Darmstadt, Germany). One slide was sent to the local Department of Pathology for hematoxilineosin (H&E) staining and cytopathologic interpretation"

Positivity threshold: quotes: "generally accepted cytopathologic criteria" "Insufficient, tumor cell negative, suspicious for tumor cells, tumor cells positive" "grouped as 'negative' in cases with benign changes or with the finding of mild dysplastic epithelial cells (SIN 1) only, and as 'positive' if cells with moderate or severe dysplasia (SIN 2, SIN 3) or malignant tumor cells were present"

Sequence of tests: index followed by reference standard

Training or calibration of clinicians: quotes: "two experienced pathologists" "In cases of false-negative or contradictory results for histology, cytology and DNA-ICM, respective slides were re-assessed by three experienced persons (BG, BS, KPW). If there was an disagreement, the respective slides were reviewed together until a mutual agreement was obtained"

Blinding of examiners: quote: "All investigators were blinded against the results of the complementary method and against the histological results which served as the gold standard"

Multiple tests: yes

Method of site selection: quote: "From each lesion"

Conflict of interests: the authors declare that they have no conflict of interest. Unfunded study

Index test 2 - brush cytology plus DNA-image cytometry

Category: cytology - cytobrush plus DNA-image cytometry

Description: quotes: "From each lesion, four slides from a separate brush biopsy were taken without prior mucosal disinfection. A proprietary brush (Cytobrush Plus GT; Medscan, Malmo, Sweden) was used to obtain epithelial cells. The small brush was rotated for 10 revolutions over an approximately 10 x 10 mm (100 mm2) area, resulting in most cases in petechial bleeding points. Immediately after harvesting, the cells were transferred onto four pre-charged glass microscope slide (SuperFrost Plus, Menzel, Braunschweig, Germany) in a rotating manner and fixed by ethyl alcohol (Merckofix, Merck, Darmstadt, Germany)" "For each specimen, 250–300 randomly chosen cells of the suspicious cell population were measured. The mean DNA content of > 30 non-neoplastic nuclei (lymphocytes) was taken as internal reference. Coefficients of variation of reference cells were below 5%. Reference extinction was normalized to the 2c (2-fold DNA content)-value of the reference cells. The standards and guidelines of the European Society for Analytical Cellular Pathology (ESACP) for DNA-ICM were followed"

Positivity threshold: quotes: "a slide was classified as DNA-diploid, if only one DNA stemline (STL) between 1.80 c and 2.20 c was detected. DNA-polyploidy was assumed, if there were DNASTLs between 1.80 c and 2.20 c and between 3.60 c and 4.40 c. A lesion was characterized as DNA-aneuploid, if an abnormal DNASTL < 1.80 c and > 2.20 c or < 3.60 c and > 4.40 c and/or 9 c exceeding events were detected" "For DNA-ICM, a case was classified as 'negative' if DNA-aneuploidy could not be proven, and as 'positive' for the finding of DNA-aneuploidy"

Sequence of tests: index followed by reference standard

Kämmerer 2013 *(Continued)*

	Training or calibration of clinicians: quotes: "All DNA-ICM measurements were conducted in duplicates by two investigators (SM, KPW) with experience in the method" "In cases of false-negative or contradictory results for histology, cytology and DNA-ICM, respective slides were re-assessed by three experienced persons (BG, BS, KPW). If there was an disagreement, the respective slides were reviewed together until a mutual agree- ment was obtained" Blinding of examiners: quote: "DNA-ICM was conducted without knowledge of the nature of the lesion or the histopathological grading" Multiple tests: yes Method of site selection: quote: "From each lesion" Conflict of interests: the authors declare that they have no conflict of interest. Unfunded study
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "After harvesting cells, a scalpel incision biopsy (approximately 0.1 ??? 0.1 cm) was taken from exact the same area. The obtained tissue was formalin-fixed, paraffin-embedded and diagnosed at the Department of Pathology" Positivity threshold: quotes: "The tissue was classified as normal (optionally with inflammatory alterations), dysplastic (subdivided in squamous intraepithelial neoplasia (SIN) 1 and > 120), or invasive carcinoma" "The histological diagnoses of the scalpel biopsies were grouped as 'negative' in cases with benign changes or with the finding of mild dysplastic epithelial cells (SIN 1) only, and as 'positive' if cells with moderate or severe dysplasia (SIN 2, SIN 3) or malignant tumor cells were present" Sequence of tests: index test followed by reference standard. Training or calibration of pathologists: quote: "two independent experienced pathologists" Blinding of examiners: quote: "two independent experienced pathologists" Multiple tests: no Method of site selection: quote: "After harvesting cells, a scalpel incision biopsy (approximately 0.1 x 0.1 cm) was taken from exact the same area. The obtained tissue was formalin-fixed, paraffin-embedded and diagnosed at the Department of Pathology" Target condition: moderate or severe dysplasia, carcinoma
Flow and timing	Patients receiving index tests but not reference test: 0 Patients receiving reference test but not index tests: 0 Time interval: not reported but assumed at the same appointment Patients receiving both index and reference test but excluded from analysis: 1 (insufficient cell yield for DNA-IC)
Comparative	
User defined 1	
Notes	Multiple tests but considered independent. Quote: "All investigators were blinded against the results of the complementary method and against the histological results which served as the gold standard." Authors also report results for the "combined" accuracy of the tests, the combination of cytology and DNA-ICM was considered as positive if either 1 or both of the tests showed positive results. In line with the remit of the review, the combined results were not used as they did not meet the inclusion criteria 'combinations of tests would be included if 1 test informed the use of the second adjunct.' In this case, brush cytology did not inform DNA-IC so the results for the combined test were not included in the review. However, the results for brush cytology and DNA-IC alone were used. Results in meta-analysis are those for all cases

Kämmerer 2013 *(Continued)*

Methodological quality

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

Kämmerer 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	
Could the patient flow have introduced bias?		Low risk

Kaur 2016

Study characteristics	
Patient Sampling	Method of patient selection: quote: "100 consecutive patients presenting with clinically suspicious lesions of the oral cavity"
Patient characteristics and setting	Age: mean age 50.4 years (SD 15.2), range 19 to 90 years Sex: 78 male, 22 female SES: not reported Ethnicity: not reported Stated risk factors: 72 smokers, 19 tobacco chewing, 12 betel nut consumers, 10 alcohol Number of patients/legions: 100/100 Lesion site: oral cavity Severity: suspicious for malignancy, malignant Country: India Type of facility: not reported Prevalence: 48/96
Index tests	Category: cytology - brush, DNA-IC Description: quotes: "The cytology material was collected with the help of a toothbrush. The brush was rotated under slight pressure several times on the suspicious looking area without local anesthesia. Brush was then immediately rolled on glass slides. A minimum of four slides were made per case. The smears were processed for cytologic examination and image analysis. For routine cytology, the smears were stained with MayGrunwald giemsa (May–Grunwald–Giemsa), hematoxylin and eosin (H & E), and Papanicolaou stain (Pap)" "Air-dried smears were stained by Feulgen method.... If the peak was narrow and the value was between 1.8 c and 2.2 c, it was taken to represent a diploid DNA content. If the peak fell less than 1.8 c or over 2.2 c, it was considered as aneuploid DNA content. Multiple peaks with distribution outside diploid range was also taken as aneuploid" Positivity threshold: suspicious for malignancy, malignant Sequence of tests: index followed by reference Training or calibration of clinicians: not reported Blinding of examiners: not reported Multiple tests: yes Method of site selection: collected by the pathologist from the most suspicious region of the lesion without prior disinfection Conflict of interests: no conflict of interest statement

Kaur 2016 *(Continued)*

Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: surgical biopsy Positivity threshold: dysplasia, malignancy Sequence of tests: index then reference Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: not reported Target condition: dysplasia, malignancy
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: reference test biopsy immediately followed the index test, quote: "Following the sampling for cytologic analysis, surgical biopsy of the same site was performed in all the cases concomitantly" Patients receiving both index and reference test but excluded from analysis: sampling considered inadequate for 4 cases
Comparative	
User defined 1	
Notes	The manuscript reports the results for standard brush cytology, DNA-IC and the results of the test combined. In line with the remit of the review, the combined results were not used as they did not meet the inclusion criteria 'combinations of tests would be included if 1 test informed the use of the second adjunct.' In this case, brush cytology did not inform DNA-IC so the results for the combined test were not included. However, the results for brush cytology and DNA-IC alone were used

Methodological quality

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

Kaur 2016 *(Continued)*

knowledge of the results of the index tests?			
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	

Koch 2011a

Study characteristics	
Patient Sampling	Method of patient selection: quote: "All patients attended the Maxillofacial Surgery Clinic at the University Hospital in Mainz, Germany and were examined between September 2005 and December 2007"
Patient characteristics and setting	Age: 62.8 +/- 18.3 years Sex: approximately 2:1 ratio SES: not reported Ethnicity: unclear Stated risk factors: not reported Number of patients/lesions: 135/182 Lesion site: oral cavity Severity: quote: "clinically diagnosed as squamous cell carcinoma (SCC) or suspicious epithelial lesions" Country: Germany Type of facility: secondary Prevalence: 113/182

Koch 2011a *(Continued)*

Index tests	Category: cytology - cytobrush Description: quote: "The cytologic samples were obtained using the Cytobrush Plus GT. The brushes were obtained from the lesions in a rotating manner without local anaesthetic so that petechial bleeding points occurred in the majority of cases" Positivity threshold: quote: "subgroups of benign hyperplasia or hyperkeratosis, mild, moderate or severe dysplasia, and SCC were used to classify the cytologic diagnoses" Sequence of tests: index followed by reference Training or calibration of clinicians: unclear Blinding of examiners: quote: "All cytologic smears were examined blindly, i.e. without information on the patient or the result of the histological examination, by an independent, experienced cytopathologist" Multiple tests: no Method of site selection: most suspicious region of the lesion was sampled Conflict of interests: not reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: scalpel, quote: "Following the sampling for cytologic analysis, all oral lesions were examined by taking a conventional biopsy by excision" Positivity threshold: 2 classification systems: SCC and carcinoma in situ; high risk lesions of SIN II/III and SCC Sequence of tests: index followed by reference Training or calibration of pathologists: unclear but 2 pathologists for each sample - 3 pathologists in total Blinding of examiners: analysed by different specialists Multiple tests: no Method of site selection: most suspicious region of the lesion was sampled Target condition: dysplasia and carcinoma
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: same time Patients receiving both index and reference test but excluded from analysis: 4
Comparative	
User defined 1	
Notes	Low attrition

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Koch 2011a *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Koch 2011b

Study characteristics	
Patient Sampling	Method of patient selection: quote: "78 patients participating in the study attended the outpatient clinic of the Oral and Maxillofacial Surgery clinic of the Mainz University Medical Centre and suffered from suspicious oral mucosal lesions"

Koch 2011b *(Continued)*

Patient characteristics and setting	Age: mean 61.7 years Sex: 59% male, 41% female SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/lesions: 78/78 Lesion site: cheek 26, gingival 22, floor of the mouth 7, sulcus glossoalv. 2, tongue lower side 2, tongue dorsum 3, palate 8, arcus palatoglossus. 5, inner lips 3 Severity: 41% red, like erythroplakia (17%) or erythroleukoplakia (24%); 21% white, like leukoplakia 21% ulcerous; 17%, a speckled, including fibrin-covered lesions Country: Germany Type of facility: university medical centre Prevalence: 33/78
Index tests	Category: light-based Description: quote: "Two different investigation methods were applied: the standard examination by white light and the examination by a 400-nm wavelength light source that is supposed to trigger a green light emission (> 500 mm) in normal mucosa." Documented by digital reflex photography Positivity threshold: SCC, and dysplasia [identified] depending on 2 different autofluorescence features: 1) a black or dark green aspect, as well as red indicating dysplasia or SCC (positive). Also, a speckled, heterotopic aspect of both green and autofluorescence negative or reddish regions indicated a positive finding; 2) the presence of red mucosal autofluorescence was evaluated as a separate indicator for dysplasia or SCC (positive) Sequence of tests: index then reference Training or calibration of clinicians: not clearly discussed Blinding of examiners: 1 experienced examiner performed the visual examination, 2 examiners examined the photographs, they were blinded and independent Multiple tests: visual then light, class as 1 index test only, as an adjunct Method of site selection: not discussed Conflict of interests: none
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "a biopsy by incision was performed. Then, the biopsies were fixed with formaldehyde 4.5% and processed for light microscopy via paraffin-embedded, haematoxylin–eosin-stained slices" Positivity threshold: mucosal hyperkeratosis, lichen planus, SCC, inflammation, dysplasia, healthy mucosa Sequence of tests: index then reference Training or calibration of pathologists: performed by 1 experienced examiner

Koch 2011b *(Continued)*

	Blinding of examiners: not described Multiple tests: no Method of site selection: biopsy of photographed lesion Target condition: SCC and dysplasia
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: biopsy taken immediately after index test Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Results calculated from sensitivity and specificity in table 2

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined			Low concern

Koch 2011b *(Continued)*

by the reference standard does not match the question?		
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

Leunig 2000

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Fifty-eight patients (mean age, 57.7 y; range, 34-74 y) with a suspected squamous cell carcinoma of the oral cavity were investigated"
Patient characteristics and setting	Setting: not clearly stated Age: 57.7 years, range 37 to 74 years Sex: not discussed SES: not discussed Ethnicity: not discussed Stated risk factors: quote: "All patients had a history of smoking and of drinking alcohol" Number of patients/lesions: 58/160 Lesion site: oral cavity Severity: not discussed Country: Germany Type of facility: not discussed Prevalence: 100/160
Index tests	Category: light-based Name of test(s): imaging 5-aminolevulinic acid-induced protoporphyrin IX fluorescence Description: quote: "5-Aminolevulinic acid..... was used in a 0.4% rinsing solution. The patients performed a 15-minute continuous rinsing of the oral cavity using the 5-ALA solution. After an incubation period of 1 to 2.5 hours, fluorescence investigation was performed with the patient either awake (n = 42) or under general anaesthesia during surgery (n = 16)"

Leunig 2000 *(Continued)*

	Positivity threshold: strong, macroscopically visible red fluorescence (F++), weak (F+) or negative (F-) Sequence of tests: index then reference Training or calibration of clinicians: unclear Blinding of examiners: unclear Multiple tests: unclear Method of site selection: light-based detection Conflict of interests: not discussed
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "Biopsy specimens were taken from tumor, tumor boundaries, and normal tissue under fluorescence illumination" Positivity threshold: quote: "histological diagnoses of squamous cell carcinoma, carcinoma in situ, and severe and moderate dysplasia were classified as malignant, whereas light dysplasia and normal tissue were classified as benign" Sequence of tests: index then reference Training or calibration of pathologists: not discussed Blinding of examiners: not discussed Multiple tests: no Method of site selection: specimens taken from identified sites Target condition: quote: "histological diagnoses of squamous cell carcinoma, carcinoma in situ, and severe and moderate dysplasia were classified as malignant, whereas light dysplasia and normal tissue were classified as benign"
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: unclear Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
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		

Leunig 2000 *(Continued)*

Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	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	

Liu 2019

Study characteristics	
Patient Sampling	Method of patient selection: unreported
Patient characteristics and setting	Age: 52.44 years (SD 13.55) Sex: 98 males, 105 females SES: not reported

Liu 2019 *(Continued)*

	Ethnicity: not reported Stated risk factors: none stated Number of patients/lesions: 203 patients Lesion site: unclear (subjects with OSCC, OPMDs and other oral mucosa disease without dysplasia according to clinical and pathologic diagnosis) Severity: unclear (no specific criteria on severity) Country: China Type of facility: secondary care (university affiliated dental hospital) Prevalence: 86/203
Index tests	Category: cytology DNA-IC Description: the DNA cytometry test was conducted with the MotiCytometer system provided by a diagnostic facility company from Xiamen, China. This system uses an automatic microscope and a camera to scan cells in the specimen slices. It uses normal cells in the specimen as internal standard, and performs automatic classifications and counting of cell nucleus based on parameters, among which DNA index (DI) is the key parameter Positivity threshold: standards for a positive result: 1) ≥ 10% hyperplasia, 2) 3 or more cells with DI ≥ 2.5, or 3) aneuploid peak visible Sequence of tests: index followed by reference Training or calibration of clinicians: not stated Blinding of examiners: examiners are blinded from results of the histopathologic assessment Multiple tests: no Method of site selection: not stated Conflict of interests: statement not provided. Funded by the Program for New Clinical Techniques and Therapies of Peking University School and Hospital of Stomatology (PKUSSNC-T-15A05)
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: surgical biopsy. After HE staining, 2 pathologists performed assessments according to the WHO classifications on epithelium dysplasia Positivity threshold: not clearly specified but possibly the WHO classifications on epithelium dysplasia Sequence of tests: index then reference Training or calibration of pathologists: unreported Blinding of examiners: examiners are blinded from results of DNA cytometry Multiple tests: no Method of site selection: unreported Target condition: OSCC, OPMDs and other oral mucosa disease without dysplasia
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: unreported Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	

Liu 2019 *(Continued)*

Notes	Translation provided by Fang Hua on 2 March 2020

Methodological quality

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

Ma 2014

Study characteristics	
Patient Sampling	Method of patient selection: quote: "52 subjects who sought examination and remedy of oral lesions at the Department of Oral Medicine"
Patient characteristics and setting	Age: mean age 58 years, range 19 to 84 Sex: 15 male, 37 female SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/legions: 52/52 Lesion site: oral cavity Severity: severe dysplasia or neoplasia Country: China Type of facility: secondary care Prevalence: 22/52
Index tests	Category: cytology - DNA-IC Description: quote: "To obtain smears, we used a cytobrush cell collector, which was rolled at the location of the mucosal lesion at least three times using gentle pressure. The exfoliative cells were transferred to two slides, which were immediately fixed with absolute ethanol. Nuclear DNA contents (ploidy) were measured after Feulgen staining using an automated DNA image cytometer" Positivity threshold: DNA-aneuploid Sequence of tests: index followed by reference Training or calibration of clinicians: not reported Blinding of examiners: not explicitly stated but "The smears and biopsy tissues were examined at different institutions" Multiple tests: no Method of site selection: suspicious lesions Conflict of interests: the authors declare that they have no conflicts of interests
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: surgical biopsy Positivity threshold: severe dysplasia or malignancy Sequence of tests: index then reference Training or calibration of pathologists: not reported Blinding of examiners: assumed based on quote "The smears and biopsy tissues were examined at different institutions" Multiple tests: no Method of site selection: suspicious lesion Target condition: severe dysplasia or malignancy
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	No information provided in report to enable re-analysis at mild+ level

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Ma 2014 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
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?	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?		Low risk	

Mashberg 1980

Study characteristics	
Patient Sampling	Method of patient selection: quote: "A thorough examination of the oral soft tissue was conducted for most patients if asymptomatic mucosal alterations were observed then patients were referred for an evaluation, then rescheduled 10-14 days later for re-evaluation and TB testing." Implied that the 14-day period is part of recruitment process and patients are not deemed part of the study if le-

Mashberg 1980 *(Continued)*

	sions do not persist. Conditions recorded squamous cell carcinoma (invasive), carcinoma in situ, atypia, or benign (hyperplasia, keratosis, inflammation, etc.)
Patient characteristics and setting	Age: not reported Sex: not reported SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/lesions: 178/235 Lesion site: oral cavity Severity: asymptomatic mucosal alterations Country: USA Type of facility: secondary - Veteran Administration Medical Center so concern over wider applicability Prevalence: 105/235
Index tests	Category: vital staining - toluidine blue Description: photographed before and after direct application of stain, twice mouthrinse with water (20 seconds each), 1 rinse with 1% acetic acid solution (20 seconds), area dried with gauze, '1% toluidine blue solution containing acetic acid and alcohol' applied with cotton swab, rinse with acetic acid solution (1 minute), rinse with water Positivity threshold: quote: "positive for malignancy if the lesion stains dark blue (royal or navy); either the entire lesion or a portion of it may stain solidly or stippled. Occasional equivocal stains are considered positive unless proven otherwise" Sequence of tests: toluidine blue then biopsy Training or calibration of clinicians: not discussed, possibly only 1 examiner Blinding of examiners: not explicitly stated but implied as biopsy results are returned for comparison with toluidine blue results Multiple tests: no Method of site selection: not described, assumed sites of lesions after examination Conflict of interests: not reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "A biopsy of the stained areas was performed" or "a biopsy of the most suspicious area as defined by our erythroplastic criteria was performed" Positivity threshold: invasive carcinoma or carcinoma in situ Sequence of tests: index then reference Training or calibration of pathologists: not discussed Blinding of examiners: not discussed Multiple tests: no

Mashberg 1980 *(Continued)*

	Method of site selection: quote: "A biopsy of the stained areas was performed" or "a biopsy of the most suspicious area as defined by our erythroplastic criteria was performed" Target condition: invasive carcinoma, carcinoma in situ, atypia, benign
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: description implied biopsy taken immediately after rinsing Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Re-analysed data to include atypia as positive (atypia defined in text as epithelial dysplasia)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard			Low concern

Mashberg 1980 *(Continued)*

does not match the question?		
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

McIntosh 2009

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "Patients presenting to an oral medicine specialist unit for assessment of an oral mucosal lesion were recruited into the study" "The only criterion for inclusion was referral for examination of an oral mucosal white lesion that was deemed to be clinically suspicious and warranted further evaluation by routine measures including definitive histopathology"
Patient characteristics and setting	Age: mean age 56.6 (range 26 to 87.2) years Sex: 27 female (mean age 57.06); 23 male (mean age 56.03) SES: not reported Ethnicity: not reported Stated risk factors: smoking habits (20/50 – 40%) and regular alcohol consumption (26/50 – 52%); 18/50 (36%) were both tobacco and alcohol consumers Number of patients/lesions: 50 patients/50 lesions Lesion site: tongue, buccal mucosa, floor of mouth, alveolar ridge, gingiva, palate, lip Quotes: "Location of primary and satellite lesions is listed in Table 2, with the majority of primary lesions occurring on the tongue and buccal mucosa" "Oral mucosal white lesions" Severity: clinically suspicious lesions, sufficient to be referred to an oral medicine specialist unit for assessment. Quote: "deemed to be clinically suspicious and warranted further evaluation by routine measures" Country: Australia Type of facility: secondary care, oral medicine specialist unit Prevalence: 9/50
Index tests	Category: light-based - Microlux Description: visual examination, with dimmed operatory lights, examiner-worn LED white-headlight, and handheld re-usable LED diffused white-light guide to illuminate irregular cells. Quotes: "clinical examination was repeated using Microlux/DL diffused light illumination kit" "examinations were undertaken first without the 1% acetic acid solution and then repeated after the recommended 60-s rinse procedure"

McIntosh 2009 *(Continued)*

	Positivity threshold: quotes: "After rinsing with the acetic acid solution, the manufacturer states that irregular cells will take on a whitish hue which will contrast with the surrounding tissues making it more obvious to the examiner" "Borders were designated either as diffuse or sharp" Sequence of tests: index followed by reference Training or calibration of clinicians: not reported Blinding of examiners: index completed prior to reference Multiple tests: yes (light with/without acetic acid) Method of site selection: not reported Conflict of interests: quote: "The authors declare that they have no financial or personal relationship with any party or product that could inappropriately influence or bias the results of this study"
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "An incisional scalpel biopsy was performed under local anaesthesia to obtain a definitive histopathological diagnosis which would serve as the gold standard. Biopsy specimens were fixed in formalin, blocked in paraffin, stained with haematoxylin and eosin, and assessed by routine histopathology by an experienced pathologist, not involved with the clinical study, using recognised measures of interpretation" Positivity threshold: quote: "A histopathological diagnosis of dysplasia was considered positive for the presence of disease in these comparisons" Sequence of tests: index test followed by reference standard Training or calibration of pathologists: unclear, however performed by experienced pathologist Blinding of examiners: quote: "not involved with the clinical study" Multiple tests: no Method of site selection: not specifically reported; however, it is implied that the biopsy selection site was intended to be the clinically suspicious lesion warranting further investigation, in addition to any satellite lesions identified during the index test Target condition: dysplasia/OSCC
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: reference performed immediately after index test Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	The analysis data for sensitivity and specificity is dysplasia/OSCC, taken from table 5

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

McIntosh 2009 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

McIntosh 2009 *(Continued)*

Could the patient flow have introduced bias?	Low risk

Mehrotra 2008

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Ninety-four patients with suspicious oral lesions from Departments of Otorhinolaryngology and Pathology, Moti Lal Nehru Medical College, Allahabad, India, were studied in a random manner"
Patient characteristics and setting	Age: range 10 to 80 years Sex: 76% male, 24% female SES: not reported Ethnicity: not reported, India Stated risk factors: 54% tobacco consumption, 35% exposure to carcinogen, 5% heavy alcohol use Number of patients/legions: 94/94, 79 provided adequate brush biopsies Lesion site: oral cavity Severity: quote: "Only lesions with an abnormal epithelial surface including erythroplakia, leukoplakia without dysplasia and oral submucous fibrosis were included" Country: India Type of facility: medical college Prevalence: 34/79
Index tests	Category: cytology - baby toothbrush Description: a hard nylon baby toothbrush, using moderate pressure, was repeatedly brushed in 1 direction over the entire lesion until pinpoint bleeding was obtained. The material from the brush was spread on to dried glass slides. The smears were fixed immediately with 100% ethanol for staining with hematoxylin and eosin and the modified Papanicolaou's method Positivity threshold: cells showing changes were categorised as malignant, quote: "enlarged nuclei, variation in nuclear size and shape (pleomorphism), nuclear borders, nucleo:cytoplasmic ratio, number of nuclei, binucleation, keratinization, tadpole forms, and hyperchromatism chromatin" Sequence of tests: index followed by reference Training or calibration of clinicians: not discussed, quote: "specimens were examined manually independently by 2 different cytopathologists," not clear how experienced these cytopathologists were Blinding of examiners: quote: "examined manually independently by 2 different cytopathologists in a double blind fashion," also blind to the reference analysis Multiple tests: no Method of site selection: not reported Conflict of interests: not discussed

Mehrotra 2008 *(Continued)*

Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "For histological diagnosis (after routine processing and paraffin embedding), several sections (3- to 4-m thickness) were cut from each case and stained with hematoxylin and eosin" Positivity threshold: WHO criteria was used to classify into grade I to III Sequence of tests: index then reference Training or calibration of pathologists: not described Blinding of examiners: quote: "All specimens were examined manually, independently by 2 different pathologists in a double-blind fashion" Multiple tests: no Method of site selection: not reported, assumed taken from the site that received the brush biopsy Target condition: dysplastic
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 15 samples were deemed inadequate for analysis so did not receive 'full' index test Time interval: reference test biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Mehrotra 2008 *(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?	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	

Mehrotra 2010

Study characteristics	
Patient Sampling	Method of patient selection: patients selected for study after detection of a clinically innocuous lesion (Class II) during routine dental care. Quote: "Patients with Class II lesions for subsequent evaluation with the light-based adjunct screening tools.... We excluded patients with Class I lesions detected with a conventional overhead examination light (and referred them for treatment) and those without any oral lesions"
Patient characteristics and setting	Age: mean age 40 years Sex: 28 female, 230 male SES: not stated Ethnicity: not stated Stated risk factors: tobacco usage 70.5% Number of patients/lesions: 258/258 split into 2 groups of 102 ViziLite+, 156 VELscope Lesion site: buccal mucosa, retromolar trigone, tongue, alveolar mucosa, gingiva, floor of mouth, palate

Mehrotra 2010 *(Continued)*

	Severity: identified as Class II prior to biopsy, so patients classified as Class I (quote: "suspicious enough to warrant a biopsy") were excluded Country: India Type of facility: quote: "outpatient department of the government-run District Hospital" Prevalence: 12/156 and 4/102, combined 116/258
Index tests	**Index test 1 - VELscope** Category: light-based Description: quote: "The clinicians performed the examinations with the VELscope and ViziLite devices according to the manufacturers' instructions" Positivity threshold: quote: "Normal mucosa - a negative VELscope finding - appears as a bright green glow, while abnormal mucosa - a positive VELscope finding - is identified by a loss of fluorescence and appears dark" Sequence of tests: quote: "depending on which screening aid was available, underwent an examination with VELscope or ViziLite. The assignments were completely random" followed by reference standard Training or calibration of clinicians: calibrated by an experienced professional from the manufacturer, no results reported Blinding of examiners: index test examiners blinded to reference test results Multiple tests: yes; conducted independently during the same session, by the same examiners Method of site selection: principle area of morphology decided by consensus after visual examination Conflict of interests: quote: "None of the authors reported any disclosures" **Index test 2 - ViziLite Plus** Category: light-based Description: quote: "The clinicians performed the examinations with the VELscope and ViziLite devices according to the manufacturers' instructions" Positivity threshold: quote: "a positive ViziLite finding - appeared aceto white. The ViziLite Plus with TBlue system also contains a toluidine blue dye, which is intended to be used only to mark lesions for follow-up examination that are positive according to the ViziLite screening" Sequence of tests: Quote: "depending on which screening aid was available, underwent an examination with VELscope or ViziLite. The assignments were completely random" followed by reference standard Training or calibration of clinicians: calibrated by an experienced professional from the manufacturer, no results reported Blinding of examiners: index test examiners blinded to reference test results Multiple tests: yes; conducted independently during the same session, by the same examiners Method of site selection: principle area of morphology decided by consensus after visual examination Conflict of interests: quote: "None of the authors reported any disclosures"
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "Using the standard scalpel technique" Positivity threshold: not specifically stated Sequence of tests: index test followed by reference standard

Mehrotra 2010 *(Continued)*

	Training or calibration of pathologists: not reported, although results interpreted by 2 examiners Blinding of examiners: index test completed before reference standard. Quote: "independently comparing pathological examination results with those obtained with these visual screening aids" Multiple tests: yes. COE; light-based detection adjunct ('randomised' to 1 of 2); biopsy Method of site selection: not specifically reported Target condition: quote: "potentially dysplastic and cancerous oral lesions"
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: assumed that reference test biopsy immediately followed the index test, but not specifically stated Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	VELscope and ViziLite plus tests reported independently so data extracted for both

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without	Yes		

Mehrotra 2010 *(Continued)*

knowledge of the results of the index tests?			
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	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	

Mehrotra 2011

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Patients who were at least 18 years of age presenting with unrelated complaints to the outpatient Department of Otorhinolaryngology, Moti Lal Nehru Medical College in Allahabad, were screened by a team of specialist and residents-in-training between July and November 2010. Patient with an oral epithelial abnormality that appeared clinically benign- minimally suspicious - and did not have an obvious etiology such as trauma or infection were prospectively enrolled"
Patient characteristics and setting	Age: mean 45.5, range 25 to 75 years Sex: 55 male, 30 female SES: not reported Ethnicity: not reported, India Stated risk factors: 37 (44%) tobacco use, 9 (11) alcohol use, both 10 (12%)

Mehrotra 2011 *(Continued)*

	Number of patients/legions: 85/85, only 79 used in results, 6 samples were inadequate Lesion site: oral cavity Severity: quote: "Patients with an oral epithelial abnormality that appeared clinically benign - minimally suspicious - and did not have an obvious etiology such as trauma or infection were prospectively enrolled" Country: India Type of facility: medical college Prevalence: 27/79
Index tests	Category: cytology Description: quote: "A specially designed brush was used to obtain a transepithelial specimen from all patients" samples were then "sent for further processing to OralCDx Laboratories® (Suffern, New York, USA)" Positivity threshold: 3 categories: negative - no epithelial abnormality; atypical - abnormal epithelial changes; positive - definitive evidence of epithelial dysplasia or carcinoma Sequence of tests: index followed by reference Training or calibration of clinicians: training yes but calibration unclear. Quote: "Training of investigators consisted of providing verbal instructions and watching video for performing the oral brush biopsy" Blinding of examiners: quote: "results were determined in a blinded fashion, independent of the scalpel biopsy" Multiple tests: no Method of site selection: not reported Conflict of interests: no competing interests
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "After the brush biopsy was performed, the same investigator performed a scalpel biopsy of the lesion and in the same location tested with the brush biopsy" Positivity threshold: not specifically stated Sequence of tests: index test followed by reference standard Training or calibration of pathologists: unclear Blinding of examiners: index test completed before reference standard Multiple tests: no Method of site selection: not specifically reported but "same part of the lesion sampled" Target condition: quote: "dysplasia or carcinoma"
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: reference test biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: 6 (inadequate brush biopsy sample)

Mehrotra 2011 *(Continued)*

Comparative			
User defined 1			
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		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Mehrotra 2011 *(Continued)*

Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	Low risk

Mojsa 2012

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Thirty consecutive patients with lesions suggestive of being premalignant identified by a conventional clinical oral examination under incandescent light were included into the study"
Patient characteristics and setting	Age: 50.3 (SD 15.7) years (range 23 to 80 years) Sex: 9 females, 21 males SES: not discussed Ethnicity: not discussed Stated risk factors: 10 current cigarette smokers, 11 prior cigarette smokers; 2 regular consumers of alcohol, 1 regular consumer of alcohol and tobacco Number of patients/lesions: 30/41 Lesion site: not discussed Severity: not discussed Country: Poland Type of facility: university medical college Prevalence: 7/41 dysplasia and SCC
Index tests	Category: vital staining plus adjunct - ViziLite Plus Description: initial visual examination, then 1% acetic acid rinse followed by evaluation under chemiluminescent light. Lesions then swabbed with 1% acetic acid, followed by application of toluidine blue and a further swab of acetic acid, and a visual examination under incandescent light to asses toluidine blue retention Positivity threshold: quotes: "chemiluminescence examination including the brightness, sharpness, surface texture, and size of the lesion using a 4-point scale (decreased, no change, slight improvement, marked improvement)" "tolonium chloride examination including the staining pattern using a 3-point scale (negative, incomplete, complete)." Not clear which level of colouration equates to negative, incomplete, or positive Sequence of tests: index then reference Training or calibration of clinicians: not discussed Blinding of examiners: not discussed Multiple tests: 1 combined test. Data taken from combined results in Table 4 Method of site selection: visual and chemiluminescent examination

Mojsa 2012 *(Continued)*

	Conflict of interests: none stated
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "incisional biopsy under local anaesthesia, and sample tissues on which biopsy was performed were submitted for histopathologic assessment" Positivity threshold: positive result: dysplasia, SCC; negative: benign keratosis, lichen planus, normal oral mucosa Sequence of tests: index then reference Training or calibration of pathologists: not described Blinding of examiners: not described Multiple tests: no Method of site selection: visual examination of index test Target condition: dysplasia or squamous cell carcinoma
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Table 4 used for data

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Mojsa 2012 *(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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	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	

Nagaraju 2010

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "The study group consisted of 60 subjects of both the sexes, 30 subjects with clinically suspicious premalignant lesions and 30 subjects with clinically suspicious malignant lesions. Subjects who fulfilled the following criteria were selected for the study: leukoplakia, speckled leukoplakia, erosive lichen planus, oral malignancy" "A provisional diagnosis of leukoplakia, speckled leukoplakia, erosive lichen planus (premalignant lesions) and oral malignancies were made on basis of clinical examination"
Patient characteristics and setting	Age: not reported Sex: not reported SES: not reported Ethnicity: not reported (India) Risk factors: not stated Number of patients/lesions: 30/30 with clinically suspicious premalignant lesions and 30/30 with clinically suspicious malignant lesions Lesion site: oral cavity Severity: premalignant lesions (degree of dysplasia), malignant lesions (degree of differentiation)

Nagaraju 2010 *(Continued)*

	Country: India Type of facility: Department of Oral Medicine and Radiology, Bapuji Dental College and Hospital, Davangere, Karnataka Prevalence: 55/60
Index tests	Category: vital staining - toluidine blue combined with Lugol's iodine Description: quote: "The subjects comfortably seated in the dental chair were examined following the methods described by Kerr et al" Positivity threshold: either or both of the tests staining positive Toluidine blue: quote: "Dark blue stain was considered as positive for lesions suspicious of malignancy, light blue retention was considered as positive for premalignant lesions unless proved otherwise by biopsy and the lesions without any retention of stain were considered as negative" Lugol's iodine: quote: "Interpretation of the Lugol's iodine stain - brown stain was considered as positive for lesions while lesions without any retention of stain were considered as negative" Sequence of tests: toluidine blue staining followed by Lugol's iodine followed by reference standard Training or calibration of clinicians: no training reported Blinding of examiners: index test completed before reference standard Multiple tests: no but a combination of 2 tests Method of site selection: quote: "Biopsy site was selected on the basis of clinical appearance and dye retention and in the sites where no retention of the stain occurred, clinical judgment directed the biopsy" Conflict of interests: not stated
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "Histopathological grading for premalignant lesions was performed as per the pathologic features suggested by Axell et al" Positivity threshold: quote: "grouped on the basis of degree of dysplasia into those with no dysplasia, mild dysplasia, moderate dysplasia and severe dysplasia. Oral malignancies were graded into well-differentiated (Grade I), moderately differentiated (Grade II) and poorly differentiated (Grade III) squamous cell carcinoma (SCC)" Sequence of tests: index test followed by reference standard Training or calibration of pathologists: not stated Blinding of examiners: not reported Multiple tests: no Method of site selection: quote: "Biopsy site was selected on the basis of clinical appearance and dye retention and in the sites where no retention of the stain occurred, clinical judgment directed the biopsy" Target condition: oral cancer and precancer (dysplasia)
Flow and timing	Patients receiving index tests but not reference test: 0 Patients receiving reference test but not index tests: 0 Time interval: biopsy based on dye retention assume within acceptable interval Patients receiving both index and reference test but excluded from analysis: 0

Nagaraju 2010 *(Continued)*

Comparative			
User defined 1			
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		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index	Yes		

Nagaraju 2010 *(Continued)*

test and reference standard?		
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

Nanayakkara 2016

Study characteristics	
Patient Sampling	Method of patient selection: patients diagnosed with oral leukoplakia lesions or suspicious oral malignancy
Patient characteristics and setting	Age: mode 51 to 65 years Sex: 149 male, 43 female SES: not reported Ethnicity: not reported Stated risk factors: 70 betel chewing, 1 smoking Number of patients/legions: 192/192 Lesion site: oral mucosa Severity: oral leukoplakia, suspicious oral malignancy Country: Sri Lanka Type of facility: not reported Prevalence: 167/181
Index tests	Category: cytology - brush and metal spatula Description: quotes: "Two cytological smears were taken from each lesion using a cytobrush plus cell collector and a metal spatula. Complete randomization was followed in deciding which technique to be used first in order to avoid the sequencing bias during the collection of samples" "In the spatula technique, a metal spatula moistened with normal saline was used to scrape the surface of the lesion to obtain a cytological smear. In the cytobrush technique, a disposable cytobrush was rolled over the lesion 5–10 times while applying firm pressure. The entire surface was scraped in small lesions, while fissured/reddish areas were scraped in heavily keratinized lesions. The scrapings were spread on an APES (3 aminopropyltriethoxysilane)-coated glass slide from one end to the other along the long axis of the slide. Smears were immediately fixed in 95% ethyl alcohol for 1 hour. Then, both smears were stained with modified Papanicolaou staining" Positivity threshold: dyskaryosis, malignancy Sequence of tests: index followed by reference Training or calibration of clinicians: yes Blinding of examiners: not reported Multiple tests: yes Method of site selection: oral lesions Conflict of interests: authors declare no conflicts of interest
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: incisional or excisional biopsy. Biopsy specimens were immediately fixed in 10% formaline and embedded in wax, and 4-lm-thick sections were cut using a rotary microtome. After staining with haematoxylin and eosin, sections were examined under a light microscope Positivity threshold: dysplasia, carcinoma in situ, squamous cell carcinoma Sequence of tests: index then reference

Nanayakkara 2016 *(Continued)*

	Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: oral lesions Target condition: dysplasia, carcinoma
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: reference test biopsy followed the index test Patients receiving both index and reference test but excluded from analysis: 11 inadequate cell samples from spatula and brush methods
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 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			Low concern

Nanayakkara 2016 *(Continued)*

does not match the question?		
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

Navone 2004

Study characteristics	
Patient Sampling	Method of patient selection: patients with lesions clinically identified as suspicious for carcinoma or dysplasia
Patient characteristics and setting	Age: 68.9 (14.33) years Sex: 49% female, 51% male SES: not reported Ethnicity: white Risk factors: not stated Number of patients/lesions: 78/78 Lesion site: not reported Severity: not clear Country: Italy Type of facility: oral pathology service of university hospital Prevalence: 45/78
Index tests	Category: cytology - cytobrush Description: the first 14 cases cytobrush was employed, for the other cases a wooden spatula. The lesion was scraped with the instrument, avoiding bleeding. Cells were spread on a glass slide, fixed and stained by Papanicolaou method Positivity threshold: not reported Sequence of tests: not reported Training or calibration of clinicians: not reported Blinding of examiners: not reported

Navone 2004 *(Continued)*

	Multiple tests: no Method of site selection: not reported Conflict of interests: not reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: the same day of cytology, 1 or more samples were taken after vital staining of the lesion with toluidine blue Positivity threshold: not reported Sequence of tests: not reported Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: some cases underwent multiple biopsy Method of site selection: toluidine blue Target condition: dysplasia and carcinoma
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: same day Patients receiving both index and reference test but excluded from analysis: 11
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Navone 2004 *(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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
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	

Navone 2008

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Patients with oral PMLs, referred to the Oral Medicine Section of the University of Turin, Italy entered this study. Patients with clinical features suggestive of carcinoma were not excluded"
Patient characteristics and setting	Age: not reported Sex: not reported SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/legions: 164/164 6 micro-biopsies did not provide adequate specimens Lesion site: oral cavity Severity: PMLs Country: Italy Type of facility: oral medicine section of the university hospital Prevalence: 73/158
Index tests	Category: cytology - curette Description: quote: "Patients first underwent a scraping from the whole surface of the oral lesion to be examined and care was taken to cause slight bleeding, so as to ensure that the basal layers of the epithelium had been sampled. This sampling was carried out using a disposable dermatological curette (Acu-Dispo

Navone 2008 *(Continued)*

	Curette, Acuderm Inc., Ft. Lauderdale, FL, USA) and the material was placed into the Thin Prep (Cytic Corporation, Marlborough, MA, USA) vial" Positivity threshold: dysplasia, carcinoma or negative, quote: "The diagnosis of dysplasia or carcinoma was based on recognized WHO criteria. The diagnosis was recorded as either negative or positive for the presence of neoplasia or dysplasia, whatever the grade" Sequence of tests: curette immediately followed by scalpel, both slides assessed simultaneously Training or calibration of clinicians: 3 pathologists participating in the study to avoid inter-examiner and intra-examiner, training/calibration not discussed Blinding of examiners: no, quote: "micro-biopsy and scalpel biopsy histological slides were assessed simultaneously by the same pathologists" Multiple tests: no Method of site selection: not reported Conflict of interests: quote: "This study has been supported in part by MURST ex-60% Universita di Torino, Ricerca Finalizzata Regione Piemonte and by a grant of Compagnia di San Paolo – Programma Oncologia, Torino, Italy"
Target condition and reference standard(s)	Category: scalpel biopsy with histopathologic assessment Description: quote: "Immediately after the curette sampling, scalpel biopsies were performed according to the following procedure. Local anaesthesia was used, taking care not to infiltrate the lesion itself to avoid artefacts; then one or more representative samples with a diameter of at least 6 mm were taken, focusing on ulcerated, red or verrucous areas where present together with a surrounding area of normal tissue, using either a scalpel or a punch" Positivity threshold: quote: "The diagnosis of dysplasia or carcinoma was based on recognized WHO criteria" Sequence of tests: index then reference for cell collection, analysis conducted simultaneously Training or calibration of pathologists: not described Blinding of examiners: not blinded Multiple tests: no Method of site selection: not reported Target condition: PMLs - dysplasia and carcinoma
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 6 samples from curette biopsy deemed inadequate, quote: "The micro-biopsies were defined adequate when a representative (at least 100 not superficial cells) strip of epithelium was present, whilst, when only horny material (anucleated cells) was present, they were defined as inadequate" Time interval: reference test biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Concern over use of: quote "The **small tissue fragments** present in the Thin Prep vial were placed into buffered formalin, to be processed histologically as micro-biopsies", while the remaining cytology specimen was stored up to be used for other investigations (not reported in this study)

Navone 2008 *(Continued)*

Contradiction between results tables and data in text. Data used for analysis from table 2 which agrees to sensitivity and specificity presented in results

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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		

Navone 2008 *(Continued)*

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

Ng 2012

Study characteristics	
Patient Sampling	Method of patient selection: quote: "retrospective chart review of a consecutive selection of patients who had both a biopsy and a concurrent QC assessment from 2008 to 2010"
Patient characteristics and setting	Age: median age 58 years Sex: 82 male, 89 female SES: not reported Ethnicity: not reported Stated risk factors: 87 smokers, 84 non-smokers Number of patients/legions: 171 Lesion site: oral mucosa Severity: PMDs and OSCC Country: British Columbia, Canada Type of facility: community referral-based oral medicine clinic Prevalence: 28/171
Index tests	Category: cytology - Oral Advance Description: QC samples were collected using the Oral Advance kit, using a flexible cytology brush. Image analysis followed by visual cytomorphologic analysis Positivity threshold: high-grade epithelial dysplasia and SCC. Quote: "A QC-positive finding is rendered when there is presence of DNA content abnormality confirmed with the cytopathologist's visual interpretation." High risk PMD and SCC Sequence of tests: index followed by reference Training or calibration of clinicians: quote: "Visual assessment of nuclear abnormality was performed according to the guidelines pertained to nuclear morphology, namely the presence of anisonucleosis and nuclear pleomorphism." No information on calibration reported Blinding of examiners: quote: "mucosal biopsy specimens were obtained from the same lesions and sent to 2 different pathology laboratories for independent evaluation in a blinded fashion" Multiple tests: no Method of site selection: quote: "multiple locations of the suspicious lesion 10 times in one direction"

Ng 2012 *(Continued)*

	Conflict of interests: 1 of the authors is a general pathology consultant at Perceptronix Medical, the company offering the Oral Advance index test
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: surgical biopsy, quote: "Oral biopsies were then taken from the same lesion in a conventional surgical fashion" Positivity threshold: quote: "The histopathologic diagnosis was classified into 4 groups according to the presence and the degree of epithelial dysplasia, as summarized in Table II: benign, low-risk PMD, high-risk PMD, and SCC" Sequence of tests: index then reference Training or calibration of pathologists: not discussed Blinding of examiners: yes, quotes: "Non-affiliated hospital based (Oral Biopsy Service, Vancouver, BC, Canada) oral and maxillofacial pathologists carried out tissue pathologic interpretation per World Health Organization guidelines" and "The oral pathologists were blinded from any QC-related information" Multiple tests: no Method of site selection: oral biopsy taken from the same site as exfoliative cytology Target condition: severe dysplasias and carcinoma in situ
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: reference test biopsy immediately followed the index test, quote: "concurrent QC assessment" Patients receiving both index and reference test but excluded from analysis: 28 cases were selected as negative QC control samples (14 cases of fibroma, 14 of squamous papilloma)
Comparative	
User defined 1	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and set-			Low concern

Ng 2012 *(Continued)*

ting do not match the review question?			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Onizawa 1999

Study characteristics	
Patient Sampling	Method of patient selection: participants had been referred for examination and treatment of oral lesions to Division of Oral and Maxillofacial Surgery, University Hospital of Tsukuba Hospital, Japan
Patient characteristics and setting	Age: mean 60, range 23 to 92 years Sex: 77 male, 53 female SES: not reported Ethnicity: Japanese Stated risk factors: not reported

Onizawa 1999 *(Continued)*

	Number of patients/legions: 130 Lesion site: oral cavity Severity: unclear Country: Japan Type of facility: secondary Prevalence: 86/130
Index tests	Category: light-based Description: fluorescence photography with ultraviolet flash Positivity threshold: quote: "The autofluorescence of the lesions was judged according to the intensity of fluorescence depicted in the film. Final judgement was subject to agreement by two or more examiners" Sequence of tests: index followed by reference Training or calibration of clinicians: unclear Blinding of examiners: index completed prior to reference Multiple tests: no Method of site selection: unclear Conflict of interests: not reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: biopsy method unclear Positivity threshold: unclear Sequence of tests: index followed by reference Training or calibration of pathologists: unclear Blinding of examiners: index was conducted independent of reference test Multiple tests: not applicable Method of site selection: unclear Target condition: carcinoma and dysplasia
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not discussed Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Data extractable for oral cancer as target condition

Onizawa 1999 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		Unclear risk	

Onofre 2001

Study characteristics

Patient Sampling	Method of patient selection: quotes: "Fifty patients with PMELs and superficial oral ulcerations suggestive of malignancy were selected from those treated at the Oral Medicine Service, Faculty of Dentistry, Araraquara, Brazil from August 1993 to May 1995 (n = 1957)" "Not included in this study were

Onofre 2001 *(Continued)*

	patients who refused to be submitted to biopsy (n = 21), those who abandoned treatment, or those who had clinically obvious invasive carcinomas or lesions without risk or suspicion of malignancy"
Patient characteristics and setting	Age: mean age 55.2 years, SD 13.4 years Sex: 22 females, 28 males SES: not reported Ethnicity: 45 white, 2 black, 2 mixed race, 1 Asian Stated risk factors: 34/50 smokers, 13/50 regular alcohol drinkers, 13 occupational regular exposure to the sun Number of patients/legions: 50/50 Lesion site: oral cavity Severity: quote: "PMELs and superficial oral ulcerations suggestive of malignancy" Country: Brazil Type of facility: quote: "hospital based sample" Prevalence: 13/50
Index tests	Category: vital staining - toluidine blue Description: quote: "All lesions were submitted to staining with an aqueous solution of 1% toluidine blue." Followed recommendations of Mashberg 1980 Positivity threshold: followed recommendations of Mashberg 1980: 1) "inadequate cell count" 2) "'negative" 3) "atypical epithelial cells" 4) "positive for dysplasia or OSCC" Atypical and positive results recorded as positive; inadequate results excluded Sequence of tests: index test followed by reference standard Training or calibration of clinicians: quote: "The clinical diagnosis and staining results were determined by 2 examiners, previously calibrated, who were specialists in oral medicine" Blinding of examiners: not explicitly reported but index test was performed prior to reference standard Multiple tests: no Method of site selection: quote: "sites were selected on the basis of the clinical appearance of the lesion and the staining result" Conflict of interests: quote: "We are indebted to the Mario A.S. Paino Laboratory of Clinical Pathology"
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: biopsy and histologic analysis Positivity threshold: any grade of dysplasia (in OL and OLP) and OSCC Sequence of tests: index test preceded reference standard Training or calibration of pathologists: quote: "Previously calibrated pathologist" Blinding of examiners: quote: "pathologist was not informed of the staining result" Multiple tests: no

Onofre 2001 *(Continued)*

	Method of site selection: quote: "The biopsy sites were selected on the basis of the clinical appearance of the lesion and the staining result. Areas retaining stain were biopsied. In sites where no retention of stain occurred, clinical judgment directed the biopsy" Target condition: squamous cell carcinomas, epithelial dysplasia, keratosis, lichen planus
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Clinical interpretation of cut-off required

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 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			Low concern

Onofre 2001 *(Continued)*

by the reference standard does not match the question?		
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?		Low risk

Petruzzi 2014

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "Patients who had a history of oral lesions or were at high risk for an oral lesion were identified and asked to participate" "Patients who did not have a lesion identified during the conventional visual exam were excluded as well as patients with advanced and clinically obvious OSCC"
Patient characteristics and setting	Age: mean age 56.7 years Sex: 22 females, 27 males SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/legions: 49/56 Lesion site: oral cavity Severity: quote: "history of oral lesions or were at high risk for an oral lesion" Country: Italy Type of facility: secondary Prevalence: 30/56
Index tests	Category: light-based - VELscope Description: quote: "The excitation blue light was projected on the oral mucosa by a dichroic mirror. Through the back of the hand piece, tissue AF > 480 nm was identified as green light" Positivity threshold: quotes: "the loss of the normal tissue fluorescence was judged as a malignant or dysplastic alteration. Red or orange fluorescence was not considered as malignant according to the literature" "VELscope pictures were rated as positive or negative" Sequence of tests: index test followed by reference standard

Petruzzi 2014 *(Continued)*

	Training or calibration of clinicians: quote: "four experienced experts in oral medicine previously calibrated. The four experts involved in this study have routinely used the VELscope since 2008 and analyzed ~2000 patients prior to this study" Blinding of examiners: not reported Multiple tests: yes Method of site selection: on the basis of a conventional visual examination Conflict of interests: authors declared no conflicts of interest Category: staining - toluidine blue Description: quote: "The oral rinsing protocol was: 20 s prerinse with 30 ml of 1% acetic acid; 20 s water rinse; 20 s rinse/gargle with 10 ml of TB solution; 20 s postrinse with 30 ml of 1% acetic acid (twice); a final water rinse" Positivity threshold: quote: "Only blue royal stained lesions were considered positive" Sequence of tests: index test followed by reference standard Training or calibration of clinicians: quote: "TB lesions were rated as positive or negative by four experienced experts in oral medicine previously calibrated in pairs" Blinding of examiners: not reported Multiple tests: yes Method of site selection: on the basis of a conventional visual examination Conflict of interests: authors declared no conflicts of interest
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "All specimens were placed in 10% buffered formalin for fixation and then submitted for histopathological evaluation by a senior oral pathologist blinded to the clinical findings" Positivity threshold: quote: "positive lesions were considered: dysplasia (mild, moderate, and severe) and OSCC (in situ, microinvasive, invasive)" Sequence of tests: index test preceded reference standard Training or calibration of pathologists: quote: "senior oral pathologist" Blinding of examiners: quote: "oral pathologist blinded to the clinical findings" Multiple tests: no Method of site selection: not reported Target condition: dysplasia or carcinoma
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported but assumed same appointment Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	

Petruzzi 2014 *(Continued)*

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
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?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Petruzzi 2014 *(Continued)*

Could the patient flow have introduced bias?	Unclear risk

Rahman 2012

Study characteristics	
Patient Sampling	Method of patient selection: pamphlets were issued inviting people to a self-examination 3-day event, 849 attended, 158 had red and white lesions, only 86 consented
Patient characteristics and setting	Age: mean 43 years (12.53) Sex: 20.93% female, 79.07% male SES: not reported Ethnicity: not reported Stated risk factors: tobacco usage 86.05%, alcohol usage 40.69%. Both tobacco and alcohol usage 59.3% Number of patients/lesions: 86 Lesion site: not reported Severity: quote: "suspected of having oral premalignant lesions or oral squamous cell carcinoma" Country: India Type of facility: quote: "3 day screening camp" Prevalence: 27/86
Index tests	**Index test 1 - toluidine blue** Category: vital staining Description: the modified Mashberg technique was used to prepare the 1% toluidine blue solution. Following a rinse of water and 1% acetic acid solution, the toluidine blue solution was swabbed, a further rinse with acetic acid solution was applied Positivity threshold: a royal navy blue colour was considered positive, light blue not specified to be positive or negative Sequence of tests: toluidine blue, cytobrush, scalpel Training or calibration of clinicians: not discussed Blinding of examiners: not discussed Multiple tests: yes Method of site selection: visual examination Conflict of interests: not discussed **Index test 2 - cytobrush** Category: cytology (following toluidine blue) Description: quotes: "Visually identified lesions were then scraped using a cytobrush" "smears were….. fixed in 95% alcohol or air dried and stained with hematoxylin-eosin, Pap and PAS stains"

Rahman 2012 *(Continued)*

	Positivity threshold: footnote table 7, negative = Class I & II, atypical = Class III, IV & V Sequence of tests: toluidine blue, cytobrush, scalpel Training or calibration of clinicians: not discussed Blinding of examiners: all specimens were examined by 2 pathologists in a double-blinded fashion, any discrepancies were resolved by a third opinion Multiple tests: yes Method of site selection: visual and staining Conflict of interests: not discussed
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: the lesions were then biopsied using either scalpel or punch technique under local anaesthesia Positivity threshold: mild dysplasia classed as negative; moderate and severe classed as positive Sequence of tests: index then reference Training or calibration of pathologists: not discussed Blinding of examiners: blinded to toluidine blue, but not visual examination, unclear regarding brush biopsy Multiple tests: no Method of site selection: visual examination Target condition: SCC, CIS, dysplasia
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: all completed during the same appointment Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Study is a direct comparison of 2 independent tests, toluidine blue and cytology, to be used independently. Results reported as separate index tests with potential for incorporation bias with index test of cytology

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Rahman 2012 *(Continued)*

Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Scheer 2011

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "Oral and VELscope examinations were performed on 64 patients referred to the Department of Oral and Craniomaxillofacial Surgery to rule out invasive squamous cell carcinoma" "Patients with advanced squamous cell carcinomas were excluded" 20 patients with previous history raises concern that examiners are already aware of patients diagnosis

Scheer 2011 *(Continued)*

Patient characteristics and setting	Age: average age of 59.8 years Sex: 25 female, 39 male SES: not discussed Ethnicity: not discussed Stated risk factors: not discussed Number of patients/lesions: 64 Lesion site: floor of mouth, palate, alveolar process, tongue, buccal mucosa Severity: patients with advanced SCC were excluded Country: Germany Type of facility: Department of Oral and Craniomaxillofacial Surgery Prevalence: 12/64
Index tests	Category: light-based - VELscope Description: quotes: "For evaluation of the suspicious lesions, the room was shaded and the hand piece was covered with a lens cover" "Through the back of the hand piece, tissue autofluorescence above 480 nm could be identified as green light" and photographs be taken to record the outcome Positivity threshold: quote:"the complete loss of the normal tissue fluorescence (fluorescence visualization loss [FVL]) was rated as malignant or dysplastic alteration. Red or orange fluorescence was not considered as malignant" Sequence of tests: index then reference Training or calibration of clinicians: not reported although pictures were rated by an oral and maxillofacial surgeon who had experience with this equipment since August 2007 Blinding of examiners: quote: "one experienced oral and maxillofacial surgeon without the knowledge of the histologic result rated all examinations" Multiple tests: no Method of site selection: not reported Conflict of interests: none reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "All lesions underwent histopathological evaluation by a trained pathologist" Positivity threshold: positive results: SIN, SCC Sequence of tests: index then reference Training or calibration of pathologists: quote: "All lesions underwent histopathological evaluation by a trained pathologist after examination with the VELscope" Blinding of examiners: not discussed Multiple tests: no Method of site selection: quote: "Biopsies were taken from the area of fluorescence loss" Target condition: SIN and carcinoma

Scheer 2011 *(Continued)*

Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test:0 Time interval: quote: "All lesions underwent histopathological evaluation by a trained pathologist after examination with the VELscope" Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Scheer 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	
Could the patient flow have introduced bias?		Low risk

Scheifele 2004

Study characteristics	
Patient Sampling	Method of patient selection: quote: "80 consecutive patients between July 2002 and September 2003." Inclusion criteria: quote: "(1) an OralCDx brush biopsy of a lesion with the clinical diagnosis oral leukoplakia (OL), oral lichen planus (OLP), or obvious oral squamous cell carcinoma (OSCC); and (2) a scalpel biopsy that had been performed within one month before or after the brush biopsy of the same lesion" Only those that received a scalpel biopsy were included in the study, not full spectrum of disease leading to potential sampling bias
Patient characteristics and setting	Age: mean age 58.6 years (SD 13.1 years) Sex: 33 females, 47 males SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/lesions: 80/96 Lesion site: oral cavity Severity: OralCDx brush biopsy of a lesion with the clinical diagnosis of OL, OLP, or OSCC. Included patients at high level of severity only Country: Germany Type of facility: quote: "hospital based sample" Prevalence: dysplasia or carcinoma 26/96 (lesion level)
Index tests	Category: cytology - OralCDx Description: quote: "Brush biopsies were taken according to instructions and sent to the Oral CDx centre in Germany" Positivity threshold: based on previous study (Sciubba 1999). Quotes: "1) 'inadequate cell count', 2) 'negative', 3) 'atypical epithelial cells', 4) 'positive for dysplasia or OSCC'" "Atypical and positive results recorded as positive; inadequate results excluded" Sequence of tests: quote: "scalpel biopsy that had been performed within one month before or after the brush biopsy of the same lesion"

Scheifele 2004 *(Continued)*

	Training or calibration of clinicians: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: not reported. Quote: "according to instructions" Conflict of interests: quote: "OralCDx test kits and Oral CDx analyses for this study were provided by the German Oral CDx centre.. Germany"
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: scalpel biopsy-based histology Positivity threshold: any grade of dysplasia (in OL and OLP) and OSCC Sequence of tests: quote: "scalpel biopsy that had been performed within one month before or after the brush biopsy of the same lesion" Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: not reported Target condition: dysplasia and squamous cell carcinomas
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: 1 month Patients receiving both index and reference test but excluded from analysis: 7/103, quote: "Overall, there were seven (6.8%) inadequate results"
Comparative	
User defined 1	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and set-			Low concern

Scheifele 2004 *(Continued)*

ting do not match the review question?			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
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	

Sciubba 1999

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Suspicious lesions (categorized as Class I) were analysed by use of both OralCDx and scalpel biopsy. Apparently innocuous lesions (categorized as Class II) that, in the investigators' opinion, required no further attention other than clinical follow-up were tested only by use of OralCDx. Patients with apparently innocuous lesions that produced abnormal OralCDx results, as defined below, subsequently were subjected to scalpel biopsy at the investigators' discretion"
Patient characteristics and setting	Age: mean 55 years, range 18 to 83 years Sex: male 443 (47%), female 502 (53%) SES: not reported Ethnicity: not reported

Sciubba 1999 *(Continued)*

	Stated risk factors: cigarette use, other tobacco, alcohol Number of patients/legions: 945/945 in total study, 298/298 'suspicious lesions' received both index and reference Lesion site: oral cavity, includes oropharynx Severity: quote: "intraoral lesions displaying an epithelial component" then classified into suspicious and innocuous Country: US Type of facility: dentists specialising in oral and maxillofacial pathology, oral medicine and oral surgery obtained the specimens in the course of their routine clinical practice Prevalence: 102/298 Note: although selecting those with a 'suspicious lesion' this is still an acceptable population
Index tests	Category: cytology - OralCDx Description: quotes: "the flat surface or circular border of the brush was placed against the surface of the lesion and, while firm pressure was maintained, rotated five to 10 times. Pinkness of tissue or pinpoint bleeding at the brush biopsy site was evidence of proper technique" "material collected on the brush then was transferred to the bar-coded glass slide and rapidly flooded with the fixative to avoid airdrying" "All OralCDx specimens were analyzed at OralScan Laboratories in Suffern, NY" Positivity threshold: negative: no epithelial abnormality; atypical: abnormal epithelial changes of uncertain diagnostic significance; positive: definitive cellular evidence of epithelial dysplasia or carcinoma; inadequate: incomplete transepithelial biopsy specimens (these specimens were excluded from the study) Sequence of tests: index then reference (assumed) Training or calibration of clinicians: quote: "The great majority of investigators had been trained in the oral brush biopsy and slide preparation technique at an investigators meeting, and all were provided with written instructions" Blinding of examiners: quote: "The pathologist analysing the OralCDx specimen was masked from all of the clinical and demographic data as well as histologic results" Multiple tests: no Method of site selection: not specified Conflict of interests: funded by OralScan Laboratories Inc. who produce OralCDx products
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "oral and maxillofacial pathologists at the investigators dental institutions histologically evaluated all scalpel biopsy specimens" Positivity threshold: results reported as malignant or dysplastic, benign or not performed Sequence of tests: index then reference test Training or calibration of pathologists: not discussed Blinding of examiners: not discussed Multiple tests: no Method of site selection: not described Target condition: malignant or dysplastic

Sciubba 1999 *(Continued)*

Flow and timing	Patients receiving index test but not reference test: Class I cases 0, but Class II cases 618 Patients receiving reference test but not index test: 0 Time interval: assumed reference test biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	All Class I received both index and reference test, Class II received index then reference - if positive to index. Can only look at results of Class I In data, classified positive and atypical as a positive result

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	

Sciubba 1999 *(Continued)*

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

Seijas-Naya 2012

Study characteristics	
Patient Sampling	Method of patient selection: quote: "samples obtained through OralCDx® on 24 patients who visited the Master of Oral Medicine, Oral Surgery and Implantology of the University of Santiago de Compostela, referred by the SERGAS (Servizo Galego de Saúde - Galician Public Healthcare System), between February 2009 and May 2010 who showed clinical and histological lesions that were consistent with oral leukoplakia in different clinical forms"
Patient characteristics and setting	Age: mean 62.38 years (SD 12.14) Sex: 12 male, 12 female SES: not stated Ethnicity: not stated Stated risk factors: not stated Number of patients/legions: 24/24 Lesion site: oral cavity Severity: quote: "showed clinical and histological lesions that were consistent with oral leukoplakia in different clinical forms" Country: Spain Type of facility: secondary care facility Prevalence: 11/24
Index tests	Category: brush cytology - OralCDx

Seijas-Naya 2012 *(Continued)*

	Description: quote: "brush sampling by performing 10 to 20 lateral or frontal rotations, in a representative area (never on an ulcer) until the area turns reddish or a light dotted hemorrhage; we then transferred the sample on to the pre-coded slide, placing the fixative to avoid contamination of the sample" Positivity threshold: positive for presence of dysplasia or carcinoma. All categories are atypical (cellular changes of uncertain diagnosis), positive for dysplasia or carcinoma, negative (normal cells) and inappropriate (incomplete transepithelial sample) Sequence of tests: index followed by reference Training or calibration of clinicians: no details stated for training or calibration Blinding of examiners: not stated Multiple tests: no Method of site selection: not stated Conflict of interests: not stated
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: surgical biopsy Positivity threshold: negative - no epithelial alteration, positive - dysplasia Sequence of tests: index followed by reference Training or calibration of pathologists: not stated Blinding of examiners: not stated Multiple tests: no Method of site selection: not stated Target condition: dysplasia
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: quote: "All patients had a sample taken with a surgical scalpel 3 weeks prior or after sampling with the OralCDx® kit to compare our results" Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
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		

Seijas-Naya 2012 *(Continued)*

Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?	No		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Sharwani 2006a

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Seventy-one patients with clinically suspicious oral leukoplakia took part in this study"
Patient characteristics and setting	Age: mean age 59 years, range 37 to 81 years Sex: not reported

Sharwani 2006a *(Continued)*

	SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/legions: 71/71 only 69 reported in results Lesion site: oral cavity Severity: quote: "clinically suspicious oral leukoplakia" Country: UK Type of facility: quote: "Maxillofacial Unit, University College Hospital (UCH) London" Prevalence: 31/69
Index tests	Category: light-based - 5 ALA Description: quotes: "Three hours prior to examination, topical application to the oral mucosa was performed via a rinsing solution of 0.4% 5-ALA hydrochloride" "Following examination, the images produced (Fig. 1) were analysed by the computer to identify the area of the highest signal" Positivity threshold: given in Figure 2. The ratio (red/green) was set at 1.2 or 1.3 as the threshold demarcation between normal and dysplastic. Unclear if this was decided prior to the study (see 2.2) Sequence of tests: index followed by reference standard Training or calibration of clinicians: unclear Blinding of examiners: unclear Multiple tests: no Method of site selection: not reported Conflict of interests: not reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "surgical biopsies were taken from various oral sites, where the majority originated from the tongue, buccal mucosa and floor of the mouth. These biopsies were examined histopathologically and were found to be either normal (normal, inflammatory or hyperkeratotic) or potentially malignant (mild, moderate or severe dysplastic)" Positivity threshold: results report as normal or potentially malignant (see above) Sequence of tests: index followed by reference test Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: quote: "Each of the patients was required to have 5-aminolevulinic acid in the form of mouth rinse prior to fluorescence imaging. Following this a surgical biopsy was acquired from the exact examination site" Target condition: malignant or dysplastic
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0

Sharwani 2006a *(Continued)*

	Time interval: assumed reference test biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: 2
Comparative	
User defined 1	
Notes	Badly reported. Used 1.2 threshold for results 1.3 level results: true positive 26, false negative 5, true negative 34, false positive 4

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Sharwani 2006a *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		Unclear risk

Sharwani 2006b

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Twenty-five patients, 13 males and 12 females, (mean age 52 years, range 41–67 years) with clinically suspicious oral leukoplakia took part"
Patient characteristics and setting	Age: mean age 52 years, range 41 to 67 Sex: 13 male, 12 female SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/legions: 25/25 Lesion site: oral cavity Severity: quote: "Clinically suspicious oral lesions" Country: UK Type of facility: quote: "Maxillofacial Unit, University College Hospital (UCH) London" Prevalence: 11/25
Index tests	Category: light-based - elastic scattering spectroscopy (ESS) Description: quotes: "The ESS system [...] is a prototype that was designed and built at the Los Alamos National Laboratory, USA. The system consists of a pulsed xenon-arc lamp for the light source, a PC-compatible spectrometer, which employs a linear charged coupled device (CCD) array for detection, an optical fibre (graded-index) based probe, and a laptop computer for system control and data display" "Three optical measurements were acquired from each of the suspected lesions; 1st measurement from the centre, 2nd from the periphery of the lesion and in between the two measurements the 3rd was acquired" Positivity threshold: quote: "In this clinical trial, all types of dysplasia were classified as 'malignant' whereas all other reports (normal, inflammation, and hyperkeratosis) were considered 'non-malignant or benign' changes" Sequence of tests: index followed by reference standard Training or calibration of clinicians: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: not specifically reported

Sharwani 2006b *(Continued)*

	Conflict of interests: quote: "We are grateful to 'The British Association of Oral and Maxillofacial Surgeons' who provided the funding for this study"
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "25 biopsies were taken from various oral sites, and examined histopathologically" Positivity threshold: quote: "In this clinical trial, all types of dysplasia were classified as 'malignant' whereas all other reports (normal, inflammation, and hyperkeratosis) were considered 'non-malignant or benign' changes" Sequence of tests: index followed by reference. Quotes: "ESS was used to examine the suspicious area of each of those patients prior to surgical biopsy" "Following the optical readings, a surgical biopsy was taken" Training or calibration of pathologists: training not reported; however, report indicates implied experience of pathologist due to precise description of process undertaken Blinding of examiners: not reported Multiple tests: no Method of site selection: not specifically reported, but indicated. Quotes: "ESS was used to examine the suspicious area of each of those patients prior to surgical biopsy" "The 25 biopsies were taken from various oral sites, and examined histopathologically" Target condition: malignant (including dysplasia) or carcinoma in situ
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: text implies reference test biopsy immediately followed the index test; quote: "Following the optical readings, a surgical biopsy was taken" Patients receiving both index and reference test but excluded from analysis: 10
Comparative	
User defined 1	
Notes	Unclear why 10 histologically confirmed hyperkeratotic lesions were excluded from the analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients			Low concern

Sharwani 2006b *(Continued)*

and setting do not match the review question?			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
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	

Shi 2019

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "Patients with oral white, red, mixed red and white lesions reporting to the Department of Oral Mucosal Diseases of our hospital from March 2016 to August 2018 were included in the current study" "a total of 517 consecutive patients were enrolled in the prospective diagnostic study"
Patient characteristics and setting	Age: mean age 51.9 years, range 22 to 85 years Sex: 238 male, 279 female SES: not reported Ethnicity: not reported

Shi 2019 *(Continued)*

	Stated risk factors: history of smoking 29.4%, history of alcohol intake 20.9%, patients with a history of treated cancer were excluded Number of patients/legions: 517/517 Lesion site: oral cavity Severity: quote: white, red, mixed white and red oral lesions Country: China Type of facility: hospital Prevalence: 331/517
Index tests	Category: light-based - VELscope Description: handheld, non-magnifying direct autofluorescence device Positivity threshold: quote: "the lesions were divided into two groups. Group LAF (loss of autofluorescence) included lesion that appeared dark compared to the surrounding unaltered tissue with pale green autofluorescence thus, indicating malignant or dysplastic change. Group RAF (retention of autofluorescence) included lesion that showed no change in autofluorescence when compared to the surrounding unaltered tissue thus, indicating no dysplastic change" Sequence of tests: index followed by reference standard Training or calibration of clinicians: not reported, quote: "oral medicine specialist" Blinding of examiners: not reported, but index followed by reference standard so may be assumed Multiple tests: no Method of site selection: not specifically reported Conflict of interests: the authors declare no competing interests
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "routine histopathologic examination" Positivity threshold: results presented for no dysplasia, mild dysplasia, moderate/severe dysplasia, malignancy Sequence of tests: index followed by reference Training or calibration of pathologists: not reported Blinding of examiners: quote: "by two oral pathologists who were blinded each other to the VELscope findings and were not involved with the clinical arm of the study" Multiple tests: no Method of site selection: quote: "When the lesion with retention of AF using VELscope, site selection for biopsy was based on white light; when the lesion with loss of AF, site selection was based on fluorescence of VELscope" Target condition: mild dysplasia
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported

Shi 2019 *(Continued)*

	Patients receiving both index and reference test but excluded from analysis: 0		
Comparative			
User defined 1			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 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		

Shi 2019 *(Continued)*

Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Shukla 2018

Study characteristics	
Patient Sampling	Method of patient selection: participants were recruited with clinically visible oral potentially malignant disorders and malignant pathologies
Patient characteristics and setting	Age: 21 and 60 years Sex: 37 males, 5 females SES: not reported Ethnicity: not reported Stated risk factors: none stated Number of patients/lesions: 42 participants/lesions Lesion site: oral cavity Severity: quote: "clinically visible oral potentially malignant disorders" Country: India Type of facility: secondary care Prevalence: 30/42 dysplasia or carcinoma
Index tests	Category: light-based - ViziLite, vital stain - toluidine blue Description: quotes: "examined conventionally under incandescent light and the findings were recorded" "followed by chemiluminescence examination where the patient was asked to rinse oral cavity by 1% acetic acid solution for 1 min. Room lights were dimmed, the oral cavity was examined under chemiluminescence light" "followed by examination with 1% toluidine blue where all visualized lesions were again swabbed with 1% acetic acid followed by local application of toluidine blue (Mashberg's recommendation) with a presoaked swab for 2 min" Positivity threshold: quotes: "Mucosa which appeared as decreased blue or dark was recorded as negative finding, and the mucosa appearing as improved 'aceto-white' was considered a positive finding" "Retention of toluidine blue was recorded for each lesion under standard incandescent light as positive or negative" Sequence of tests: chemiluminescence then stain, followed by reference standard Training or calibration of clinicians: not reported Blinding of examiners: not reported Multiple tests: yes Method of site selection: oral lesion Conflict of interests: authors declared no conflicts of interest

Shukla 2018 *(Continued)*

Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "paraffin-embedded specimens were cut into 3–4 µm thick sections and stained with hematoxylin-eosin" Positivity threshold: quote: "negative (acanthosis, inflammatory lesions), positive which included dysplasia (subdivided into mild dysplasia, moderate dysplasia, severe dysplasia and carcinoma in situ), oral lichen planus, oral submucous fibrosis, proliferative verrucous leukoplakia and invasive carcinoma (squamous cell carcinoma [SCC] and verrucous carcinoma)" Sequence of tests: reference follows index tests Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: lesion Target condition: dysplasia and carcinoma. Despite the positivity threshold indicated in the text, for the sensitivity and specificity results, oral lichen planus, and oral submucous fibrosis without dysplasia have been classed as negative
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: assumed reference standard biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: 2
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			

Shukla 2018 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
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	

Silverman 1984

Study characteristics	
Patient Sampling	Method of patient selection: quote: "The study group comprised 132 consecutive patients seen in the oral medicine clinic who were suspected of having oral carcinomas or precancerous (dysplastic) lesions"
Patient characteristics and setting	Age: not reported Sex: not reported SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/lesions: 132/132 Lesion site: oral cavity Severity: quote: "suspected of having oral carcinomas or precancerous (dysplastic) lesions"

Silverman 1984 *(Continued)*

	Country: USA Type of facility: secondary oral medicine clinic Prevalence: 99/132 (42 dysplastic and 57 carcinoma)
Index tests	Category: vital staining - toluidine blue Description: quote: "For staining, a 1% aqueous toluidine blue dye was applied for approximately 30 seconds, followed by a tap water rinse, and then lightly blotted with 1% acetic acid. Both solutions were applied with cotton tipped applicators. Any dye uptake was recorded by photographs" Positivity threshold: quote: "If there was dye uptake, the biopsy specimen was taken from that area; otherwise, clinical judgment guided the biopsy site" Sequence of tests: toluidine blue followed by biopsy Training or calibration of clinicians: unclear Blinding of examiners: yes. Inferred although not stated Multiple tests: no Method of site selection: quote: "If there was dye uptake, the biopsy specimen was taken from that area; otherwise, clinical judgment guided the biopsy site" Conflict of interests: not stated
Target condition and reference standard(s)	Category: biopsy Description: quote: "The biopsy specimens were fixed in 10% neutral buffered formalin and sent to the oral pathology laboratory for routine processing" Positivity threshold: categorised as: benign, dysplasia, carcinoma Sequence of tests: toluidine blue followed by biopsy Training or calibration of pathologists: not stated Blinding of examiners: not stated Multiple tests: no Method of site selection: quote: "If there was dye uptake, the biopsy specimen was taken from that area; otherwise, clinical judgment guided the biopsy site" Target condition: quote: "oral lesions suspected of being precancerous or malignant"
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: simultaneous biopsy Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Dysplasia and carcinoma extracted as positive for histology
Methodological quality	

Silverman 1984 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Singh 2015

Study characteristics

Singh 2015 *(Continued)*

Patient Sampling	Method of patient selection: quote: "selecting 50 patients with lesion in oral cavity having suspicion of malignancy"
Patient characteristics and setting	Age: mean age 49.2 years Sex: quote: "Male preponderance with male to female ratio of 8:5" SES: not reported Ethnicity: not reported Stated risk factors: quote: "Majority of the cases had multiple risk factors but maximum had tobacco consumption in some or other form" Number of patients/lesions: 50/50 Lesion site: oral cavity Severity: quote: "suspicion of malignancy" Country: India Type of facility: secondary Prevalence: 46/50
Index tests	Category: staining - toluidine blue Description: quote: "The patient was asked to rinse the mouth with 1% acetic acid for 20 s 1% toluidine blue (w/w) was then applied with the help of a swab stick and retained for 20 s. Then again rinse with 1% acetic acid for 20 s to eliminate mechanically retained stain" Positivity threshold: quote: "A dark blue (royal or navy) stain is considered positive if either the entire lesion being stained or a portion of it is stained or stippled... A light blue staining is considered doubtful. If there is no colour absorbed by the lesion, it is taken as a negative stain" Sequence of tests: index followed by reference standard Training or calibration of clinicians: not reported Blinding of examiners: index followed by reference standard Multiple tests: no Method of site selection: suspicious lesions based on clinical examination Conflict of interests: no conflict of interest statement provided
Target condition and reference standard(s)	Category: punch biopsy with histopathologic assessment Description: quote: "punch biopsy was taken from the site after applying topical anaesthetic agent (xylocaine 4 %)" Positivity threshold: not clear but appears to be severe dysplasia or malignancy Sequence of tests: reference standard follows index test Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no

Singh 2015 *(Continued)*

	Method of site selection: quote: "The biopsy site was selected on the basis of clinical appearance and dye retention and in the sites where no retention of the stain occurred, clinical judgment directed the biopsy site" Target condition: quote: "intraoral malignancies"
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: assumed reference standard biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
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		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Singh 2015 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Skandarajah 2017

Study characteristics	
Patient Sampling	Method of patient selection: patients attending the 'dental out-patient department.' After inspection by a dental surgeon, those identified as having clinically suspicious lesions were selected
Patient characteristics and setting	Age: mean age 50.6 years (SD 13), range 23 to 72 years Sex: 8 female, 24 male SES: not reported Ethnicity: not reported Stated risk factors: 7 no risk factors, 21 tobacco, 4 tobacco and alcohol drinking Number of patients/lesions: 32/32 Lesion site: oral cavity Severity: clinically suspicion lesions Country: India Type of facility: secondary Prevalence: 19/32 OSCC, 31/32 any form of dysplasia, OSCC and lymphoma
Index tests	Category: cytology - conventional and automated web-based system of digital images - cytobrush Plus GT Description: quotes: "Brush biopsy samples were immersed in 1.0 mL of SurePath cell preservative" "The cell suspension was centrifuged" "The cells were re-suspended, placed on a slide" "After drying, the cells were fixed and stained" "the prepared slide was loaded into the automated mobile microscope using a touch interface provided by a mobile app running on the iPad for image acquisition. After the images were collected, they were transmitted from the CellScope device to a web server." Samples were assessed remotely via the mobile app and conventionally Positivity threshold: "most likely positive" where cells including polynucleation, mitotic figures, anisochromatism, anisokaryosis, large nucleoli, decreased amount of cytoplasm, high nucleus to cytoplasm ratio, nuclear hypertrophy, hyperchromatism, or nuclear membrane irregularities. Otherwise deemed to be "most likely negative"

Skandarajah 2017 *(Continued)*

	Sequence of tests: 2 types of cytology assessed separately but appears to be same examiners assessing both, prior to reference standard Training or calibration of clinicians: described as expert pathologists, no calibration but diagnosis for conventional cytology was by consensus Blinding of examiners: 2 examiners assessed independently Multiple tests: yes web-based prior to conventional assessment Method of site selection: cytology Conflict of interests: part funded by Siemens
Target condition and reference standard(s)	Category: scalpel biopsy with histopathologic assessment Description: quotes: "Histological processing and examination were performed" "according to standard methods" Positivity threshold: dysplasia or carcinoma Sequence of tests: followed index test Training or calibration of pathologists: not reported Blinding of examiners: quote: "Blinded review" Multiple tests: no Method of site selection: not reported Target condition: not reported
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: assumed reference standard biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Results taken for consensus of the conventional cytology and histopathology from table 6. Data for web-based system not provided as a consensus of 2 pathologists so not extracted for the meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		

Skandarajah 2017 *(Continued)*

Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 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?			Unclear
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	

Svirsky 2002

Study characteristics	
Patient Sampling	Method of patient selection: 298 patients submitted for a scalpel biopsy that were known to have also received a brush biopsy which proved to be abnormal
Patient characteristics and setting	Age: mean 52 years, range 18 to 89 years Sex: 51% female, 49% male

Svirsky 2002 *(Continued)*

	SES: not specified Ethnicity: not specified Stated risk factors: not specified Number of patients/legions: 298 Lesion site: ventral/lateral tongue 90; palate 63; gingiva 65; buccal/alveolar mucosa 43; floor of mouth 8, unspecified/other 29 Severity: patients receiving a brush biopsy Country: US Type of facility: pathology laboratories Prevalence: 93/298
Index tests	Category: cytology - OralCDx Description: brush biopsy with computer assisted method of analysis (OralCDx) Positivity threshold: thresholds unclear Sequence of tests: index then reference Training or calibration of clinicians: not discussed Blinding of examiners: not discussed Multiple tests: no Method of site selection: not clarified, but discussion around differing results due to brush and scalpel taking sample from different parts of lesion Conflict of interests: stated no conflict of interests, declaration of some funding from OralCDx, but involvement of OralCDx laboratories is stated for retrospective analysis
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: biopsy process not described Positivity threshold: dysplasia or carcinoma Sequence of tests: index then reference Training or calibration of pathologists: not discussed, based on data from existing patients rather than allocated specifically for the study Blinding of examiners: not discussed Multiple tests: no Method of site selection: not explicitly reported Target condition: dysplasia or carcinoma
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not discussed Patients receiving both index and reference test but excluded from analysis: Unclear

Svirsky 2002 *(Continued)*

Comparative	
User defined 1	
Notes	Lack of reporting on specifics of the methods undertaken for brush and scalpel biopsies, also the patient recruitment, timing, blinding, calibration and number of assessors. Suspect this is because the study tracks those patients with a scalpel biopsy and traces them back to the index test, so the author would not know how the brush biopsy was undertaken Assuming that there was a visual examination prior to the brush biopsy

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
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		

Svirsky 2002 *(Continued)*

Were all patients included in the analysis?	Unclear
Could the patient flow have introduced bias?	Unclear risk

Trakroo 2015

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Patients with suspicious premalignant and malignant lesions were selected irrespective of age and gender"
Patient characteristics and setting	Age: range 20 to 70 years Sex: 43 male, 7 female SES: not reported Ethnicity: not reported Stated risk factors: not reported Number of patients/legions: 50/50 Lesion site: oral cavity Severity: premalignant and malignant lesions Country: India Type of facility: secondary Prevalence: 32/50
Index tests	Category: cytology - brush Description: quote: "Patients were made to rinse the oral cavity thoroughly with water, and the lesion was visualized. The brush was repeatedly brushed in one direction over entire lesion many times with moderate pressure until pinpoint bleeding was obtained, and thus obtaining epithelial cells through the full thickness of the epithelium (Figures 1 and 2). Removed cells were transferred to a glass slide by distributing the obtained material evenly over the glass surface, and smear was made. The slides were then flooded with fixative solution (95% alcohol). The cellular sample on the slide was stained with Hematoxylin and Eosin and Papanicolaou method and scanned by light microscopy" Positivity threshold: classification of Papanicolaou (1960) from Class I to Class V Sequence of tests: index then reference. Training or calibration of clinicians: not reported Blinding of examiners: index followed by reference Multiple tests: no Method of site selection: not reported Conflict of interests: stated no conflict of interests or funding

Trakroo 2015 *(Continued)*

Target condition and reference standard(s)	Category: incisional biopsy with histopathologic assessment Description: not reported Positivity threshold: quote: "The biopsy sections obtained by incisional biopsy were interpreted by an oral pathologist and based on WHO, dysplasia and squamous-cell carcinoma were graded. Grading of dysplasia was done as mild, moderate, severe and carcinoma in situ" Sequence of tests: index then reference Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: not reported Target condition: dysplasia or carcinoma
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported but assumed at the same appointment Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowl-	Unclear		

Trakroo 2015 *(Continued)*

edge of the results of the index tests?			
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Ujaoney 2012

Study characteristics	
Patient Sampling	Method of patient selection: quote: "Consecutive outpatients who visited the study centre and who clinically presented with at least one precancerous lesion were recruited into this study" Exclusions: frank malignancy, hypersensitivity to ingredients in the light examination, and any systemic disease
Patient characteristics and setting	Age: 44.4 years (17.1) Sex: female 4, male 51 SES: not discussed Ethnicity: not discussed Stated risk factors: tobacco users (29%), tobacco and lime (60%), snuff (3.6%), betel nut (63.6%), and alcohol (20%) Number of patients/lesions: 55/99 Lesion site: tongue, palate, buccal mucosa, buccal vestibule, commensural mucosa, retromolar area, labial vestibule Severity: lesions other than Class I (clinically diagnosed) Country: India Type of facility: Oral Diagnosis, Medicine and Radiology Department of the Sharad Pawar Dental College

Ujaoney 2012 *(Continued)*

	Prevalence: 17/99
Index tests	Category: vital staining plus light - ViziLite plus (reported separately) Description: quote: "combination of chemiluminescence and toluidine blue (stain) retention test" Positivity threshold: quote: "we considered a lesion to be CHTB-positive if it was both CHEM-positive and TBLU-positive; otherwise the lesion was considered to be CHTB-negative" CHTB = combined chemiluminescence and toluidine blue CHEM = chemiluminescence only TBLU = toluidine blue only Sequence of tests: chemiluminescence, then toluidine blue, then reference test Training or calibration of clinicians: not reported Blinding of examiners: not reported Multiple tests: no, class as 1 index test Method of site selection: conventional visual examination then chemiluminescence and staining Conflict of interests: no
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "collected in 10% formalin solution and processed" Positivity threshold: dysplasia (moderate and severe) or carcinoma classified as positive Sequence of tests: index then reference Training or calibration of pathologists: not discussed Blinding of examiners: quote: "Histopathologic evaluation was done by two senior Oral Pathologists blinded to the clinical findings" Multiple tests: no Method of site selection: not reported Target condition: no dysplasia, mild dysplasia, moderate dysplasia, and severe dysplasia
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: all conducted in 1 session Patients receiving both index and reference test but excluded from analysis: unclear
Comparative	
User defined 1	
Notes	Performed at lesion level. The ViziLite and toluidine blue tests are not independent as the ViziLite results were used to inform the toluidine blue site selection. Quote: "Visual examination was then repeated under standard incandescent light to identify toluidine blue retention (Figure 1C) for each previously identified lesion [from the ViziLite test] and/or any new lesions subsequently found." The study also reports the results of a combined test ViziLite Plus, which makes the toluidine blue results redundant in this context

Ujaoney 2012 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Upadhyay 2011

Study characteristics	
Patient Sampling	Method of patient selection: quote: "47 patients visiting the Dental clinics of Manipal College of Dental Sciences, Manipal, under routine OPD" Exclusion of severe cases, those included are: quote: "homogeneous leukoplakia, speckled leukoplakia, erythroplakia & erosive lichen planus," not a representative spectrum of cases
Patient characteristics and setting	Age: mean 53.83 years, range 31 to 75 years Sex: 10 female, 37 male SES: not discussed Ethnicity: not discussed Stated risk factors: not discussed Number of patients/lesions: 47 Lesion site: not discussed Severity: quote: "Clinically a provisional diagnosis of homogeneous leukoplakia, speckled leukoplakia, erythroplakia & erosive lichen planus" Country: India Type of facility: college of dental science Prevalence: 27/47
Index tests	Category: vital staining - toluidine blue Description: oral examination, water rinse for 20 seconds, 1% acetic acid for 20 seconds, rinse 5 ml of 1% toluidine blue, rinse 1% acetic acid 20 seconds, water rinse Positivity threshold: used Mashberg levels, quote: "doubtful light blue stain was considered as positive until biopsy proves the contrary" Sequence of tests: index then reference Training or calibration of clinicians: not discussed Blinding of examiners: not discussed Multiple tests: no Method of site selection: visual examination Conflict of interests: no conflict
Target condition and reference standard(s)	Category: biopsy with histopathologic examination Description: quotes: "A suitable incisional/punch biopsy was obtained on the basis of site retaining the stain" "Biopsy was also obtained from those lesions which did not retain any stain but were clinically suggestive of a PMOLs" Positivity threshold: quote: "histologically categorized as (a) Benign: hyperkeratosis, hyperplasia, & other non-malignant lesions, (b) Dysplasia: mild, moderate & severe dysplasia; and finally (c) Oral squamous cell carcinoma (OSCC)" Sequence of tests: index then reference

Upadhyay 2011 *(Continued)*

	Training or calibration of pathologists: not discussed Blinding of examiners: quote: "reviewed by two oral pathologists blinded to toluidine blue staining results" Multiple tests: no Method of site selection: based on toluidine blue staining Target condition: see positivity threshold above
Flow and timing	Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: not reported, implied to be taken during the same appointment Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Included the OCSS results in the data table

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	

Upadhyay 2011 *(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	

Warnakulasuriya 1996

Study characteristics	
Patient Sampling	Method of patient selection: quotes: "The study was conducted in August 1993 in 7 centres in Asia among a population of Sri Lankan and Pakistani subjects who consented to participate in the programme. The consultant dental surgeon in each centre approved the appropriateness of each included case" "All patients had been referred to, or had attended, the specialist centres with unconfirmed oral mucosal lesions"
Patient characteristics and setting	Age: mean age 60 years (SD 15 years) Sex: 73 male, 29 female SES: not reported Ethnicity: Sri Lankan and Pakistani Stated risk factors: 84 regular betel quid chewers (49 used tobacco in quid mixture); 28 tobacco smokers; 4 Niswar users; 13 "known to misuse alcohol" Quote: "None had previously received regular dental care" Number of patients/lesions: 102/145 only 86 received index and reference test Lesion site: oral Severity: invasive and dysplastic lesions (cf benign keratoses) Country: 7 centres in Asia Type of facility: secondary care (implied) Prevalence: 57/86 lesions
Index tests	Category: vital staining - toluidine blue

Warnakulasuriya 1996 *(Continued)*

	Description: quote: "The rinse protocol using OraScan [1% toluidine blue] was as per manufacturer's instructions... except that the study was limited to a single TB rinse per person and not repeated 14 days later" Positivity threshold: results classed as positive or negative (including equivocal) but no thresholds reported. The equivocal lesions were excluded Sequence of tests: staining followed by reference standard Training or calibration of clinicians: no training reported Blinding of examiners: index test completed before reference standard Multiple tests: no Method of site selection: toluidine blue rinse - no site selection Conflict of interests: KASSW supported by Dunhill Medical Trust. Consumables in project funded by Zila Pharmaceuticals
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "Tissues were bisected so that one portion of the biopsy could be fixed in buffered formal saline and the other in 10% alcohol; specimens were then transported to London. Following processing in wax, 6-jim sections were cut at two levels and stained with H&E and PAS, Microscopy diagnosis and, where relevant, degree of dysplasia were recorded independently by two experienced histopathologists blinded to the dye results" Positivity threshold: quote: "Dysplasias were graded as mild, moderate or severe, using the intensity of the signs specified by SMITH & PiNDBORG. When there was disagreement, concordance was reached following consultation" Sequence of tests: index test followed by reference standard Training or calibration of pathologists: not stated, but experienced histopathologist Blinding of examiners: quote: "blinded to the dye results" Multiple tests: no Method of site selection: quote: "Sites stained by the dye were charted and re-photographed. 86 clinically detected lesions, dye-retained or not, were biopsied. Ten patients had two biopsies taken from separate areas of their lesions corresponding to the dye result, one from a stain-positive area and another from a stain-negative site" Target condition: oral cancer and dysplasia
Flow and timing	Patients receiving index tests but not reference test: patient data not given, of 145 lesions 86 (or 87 in abstract) received biopsy, 59. Not reported how biopsies were selected Patients receiving reference test but not index tests: 0 Time interval: biopsy consecutive to staining Patients receiving both index and reference test but excluded from analysis: 7 toluidine blue equivocal excluded from analysis
Comparative	
User defined 1	
Notes	

Warnakulasuriya 1996 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		

Warnakulasuriya 1996 *(Continued)*

Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	High risk

Yamamoto 2017

Study characteristics	
Patient Sampling	Method of patient selection: patients referred for suspected leukoplakia or carcinoma, participant selection not reported
Patient characteristics and setting	Age: mean age 60.5 years, range 21 to 85 years Sex: 31 male, 31 female SES: not reported Ethnicity: not reported Stated risk factors: none stated Number of patients/lesions: 79 specimens obtained from 62 patients Lesion site: tongue Severity: suspected leukoplakia or carcinoma Country: Japan Type of facility: secondary care Prevalence: 64/79
Index tests	Category: light-based - VELscope (subjective); vital stain - iodine Description: - subjective method: VELscope at 15 cm - stain: quote: "washed with physiological saline solution and air-dried, and 1.5% diluted iodine was then applied. After 3 min, the area was fully rinsed and all surplus iodine liquid was removed" - objective method: VELscope camera attachment kit (doctorseyes LLC, Vista, CA, USA) was used to capture images of lesions which were "viewed from the VELscope were analyzed by Image J software (ver 1.47; National Institutes of Health, Bethesda, MD, USA), and the luminance ratio was calculated" Positivity threshold: - subjective method: quote: "green light area was defined as healthy mucosa area, and the black area of FVL was defined as lesion area" - stain: quote: "stained area was defined as the epithelial dysplasia-negative area, and the unstained area was defined as the epithelial dysplasia-positive area" - objective method: threshold determined from ROC curve results, quote: "The maximal point of addition of the sensitivity and specificity on the ROC curve was defined as the cutoff LNR value according to the Youden index" Sequence of tests: light-based followed by stain, followed by reference standard and objective test

Yamamoto 2017 *(Continued)*

	Training or calibration of clinicians: quote: "two specialists in oral and maxillofacial surgery who had 19 and 27 years of clinical experience in oral cancer treatment" completed the subjective and staining tests Blinding of examiners: not reported; suggested same examiners for conventional, subjective and stain Multiple tests: yes Method of site selection: identified lesions Conflict of interests: none reported
Target condition and reference standard(s)	Category: biopsy with histopathologic assessment Description: quote: "the lesion was resected" Positivity threshold: quote: "Histopathological diagnosis of SCC and leukoplakia with dysplasia (LwD) was defined as the epithelial dysplasia-positive group, and that of leukoplakia without dysplasia (Lw/oD) was defined as the epithelial dysplasia-negative group" Sequence of tests: following index tests Training or calibration of pathologists: not reported Blinding of examiners: not reported Multiple tests: no Method of site selection: quote: "both the blackest point identified by the subjective AVM and the iodine-unstained area" Target condition: SCC and dysplasia
Flow and timing	Patients receiving index tests but not reference test: 0 Patients receiving reference test but not index tests: 0 Time interval: Biopsy consecutive to index tests Patients receiving both index and reference test but excluded from analysis: 0
Comparative	
User defined 1	
Notes	Objective results based on the lesion to normal uptake ratio as indicated by the luminance of the healthy/lesion tissue

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		

Yamamoto 2017 *(Continued)*

Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 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?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

ALA = aminolevulinic acid; CIS = carcinoma in situ; COE = conventional oral examination; DNA-IC = DNA-image cytometry; ESS = elastic scattering spectroscopy; H&E = haematoxylin and eosin; OLP = oral lichen planus; OSCC = oral squamous cell carcinoma; PMDs = potentially malignant disorders; PMELs = potentially malignant epithelial lesions; ROC = receiver operating characteristic; SD = standard deviation; SES = socio-economic status; TB = toluidine blue; WGA-FITC = wheat germ agglutinin-fluorescein isothiocyanate.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Abram 2019	Development of a risk index, not a DTA study

Study	Reason for exclusion
Betz 2002	Inappropriate patient selection, quote: "85 patients with a histologically proven malignancy." For inclusion the index test is to be performed on patients presenting with clinically evident lesion, who are not already diagnosed
Bhoopathi 2009	Only reports on those with positive or atypical brush biopsy
Burkhardt 2010	Letter linked to Hohlwleg-Majert 2009
CTRI/2020/01/022999	Screening not diagnostic
Deuerling 2019	Retrospective analysis
Driemel 2007	Unsure of patient sampling to obtain brush biopsies or when the biopsies were undertaken. The index test and reference standard were not performed at the same time
Driemel 2007b	Unsure of patient sampling to obtain brush biopsies or when the biopsies were undertaken
Driemel 2008	Inappropriate classification of inadequate samples
Ebenezar 2012	Inappropriate population
Epstein 1995	Reference test not a scalpel biopsy
Gomez Serrano 1989	Not all patients received both index and reference test
Hedge 2006	Not all presenting with oral lesions, quote: "oral lesions or mucosal alterations suspicious of malignancy"
Hohlwleg-Majert 2009	75 investigated, 6 excluded: 1 number of cells inadequate, 1 topographical error biopsy, 3 no histology. See critical comments by Burkhardt 2010 querying histology and results. Results could not be extracted
Jayaprakash 2009	Data compromised by the inclusion of normal/control lesions
Kulapaditharon 1998	Insufficient data
Lane 2012	Data unavailable for true positives, true negatives, false positives, false negatives. Author contacted but no reply
Levine 1998	Unclear how patients were selected or clinical pathway
Li 2004	All patients recently diagnosed with oral cancer
Majumder 2006	Patients had no clinically evident lesions
Mallia 2010b	Patients had no clinically evident lesions
Maraki 2004	Could not obtain satisfactory data to calculate sensitivity and specificity
Maraki 2006	Could not obtain satisfactory data to calculate sensitivity and specificity
Navone 2007	Presentation of 'inadequate' cells made the calculation of sensitivity and specificity impossible from the data provided
Navone 2009	Data not presented to allow interpretation of sensitivity and specificity

Study	Reason for exclusion
Nieman 2008	Difficulties in clarifying the date, no response from contact author
Poate 2004	60 of the 75 who had a negative brush biopsy result did not have an incisional biopsy as the reference standard
Rana 2012	Reference test only performed on positive index test results
Reboires-Lopez 2012	No reference test
Remmerbach 2001	Reported on number of samples (smears) taken, not at patient level
Remmerbach 2003	Case-control study
Remmerbach 2004	Quote: "1328 exfoliative smears of 332 different lesions were compared with histology and/or clinical follow-ups," reported at smear/lesion not patient level
Remmerbach 2007	Reported on number of samples (smears) taken, not at patient level, same numbers of patients and lesions as 2004 study
Remmerbach 2009	Not clear that 2 x 2 tables presented actually present the combined test which is set out in the study objectives and emphasised in the discussion
Sahu 2017	Conducted on previously exfoliated cell specimens with healthy volunteers, so not appropriate sample
Sandler 1964	Diagnosis of cytology was used to determine whether biopsy would be undertaken
Schwarz 2009	66 patients without suspected oral lesions were included
Sekine 2017	The samples (including data) were retrospectively collected in clinical practice
Shklar 1970	DTA study on a highly selected group. Selection criteria was histological diagnosis of epidermoid carcinoma
Silverman 1992	Re-publication of 1984 study
Simonato 2017	Randomly selected patients with/without lesions recruited from a dental screening clinic
Sun 2019	Development of clinical risk model. Not a DTA study
Swider 1984	No data for sensitivity or specificity
Torres-Rendon 2009	Takes an archive of samples rather than using the adjunctive test in the secondary care setting
Uthoff 2018	Reference standard a conventional oral examination, no biopsy
Wang 2009	No data for sensitivity or specificity. It also uses 40 normal controls
Yang 2017	All study participants had a clinical and pathologic diagnosis of oral leukoplakia (high-grade versus low-grade epithelial dysplasia) upon enrolment in the study

DTA = diagnostic test accuracy.

Characteristics of ongoing studies *[ordered by study ID]*

CTRI/2018/02/012257

Study name	Oral cancer early detection project, a population based study (Early identification of potentially malignant and malignant lesions using clinical oral examination and VELscope on large population)
Target condition and reference standard(s)	OPMDs and OSCC
Index and comparator tests	VELscope
Starting date	1 March 2018
Contact information	MNJ Institute of Oncology and Regional Cancer Centre, Andhra Pradesh, India
Notes	CTRI/2018/02/012257

CTRI/2019/02/017623

Study name	Developing an efficient and cost effective method for screening of oral cancer in India (A pilot study to develop an efficacious oral cancer screening strategy for India)
Target condition and reference standard(s)	OPMDs and OSCC, biopsy histopathology
Index and comparator tests	Autofluorescence
Starting date	1 March 2019
Contact information	Dr B Borooah Cancer Institute, Guwahati, Assam, India
Notes	CTRI/2019/02/017623

CTRI/2019/11/022167

Study name	Mobile oral cancer screening in limited resource setting (Low-cost mobile oral cancer screening for low resource setting)
Target condition and reference standard(s)	Leukoplakia and other disturbances of oral epithelium, including tongue Malignant neoplasms of lip, oral cavity, and pharynx Biopsy
Index and comparator tests	Frontline healthcare providers using point of care diagnostic tool, artificial intelligence diagnosis
Starting date	Date of first enrolment 2 December 2019
Contact information	Dr Praveen BN, Department of Oral Medicine and Radiology, KLE Society's Institute of Dental Sciences, Bengaluru, Karnataka, India
Notes	CTRI/2019/11/022167 Unclear from the information provided in the Indian Clinical trials registry whether this study is a screening/diagnostic study

OPMDs = oral potentially malignant disorders; OSCC = oral cavity squamous cell carcinoma.

DATA

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

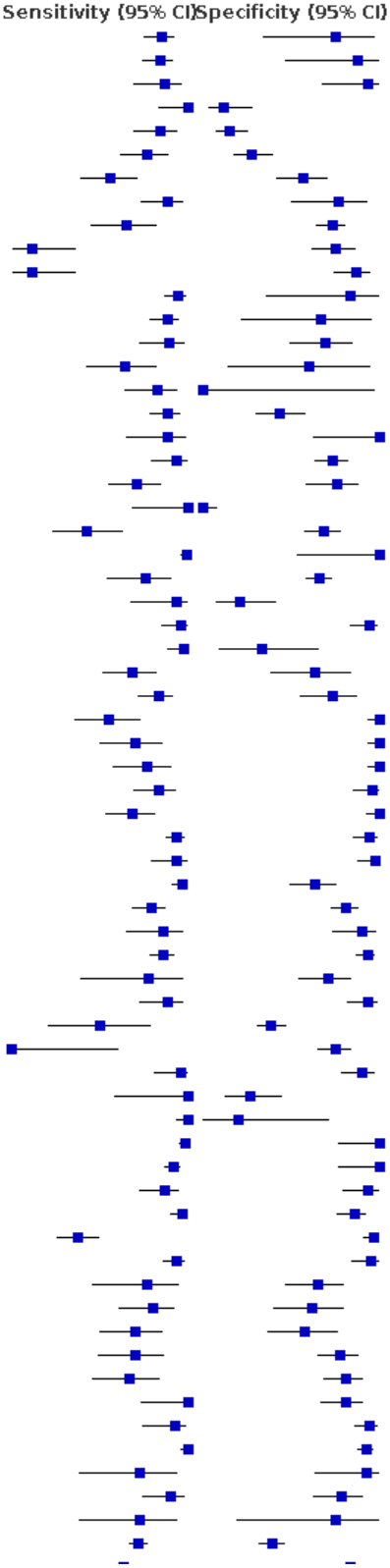
Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 All	64	7942
2 Vital staining with application method covariate	21	1780
3 Cytology with covariate	20	2892
4 Light-based with covariate	23	2587
6 Vital staining plus adjunct (light or cytology)	9	683

Test 1. All

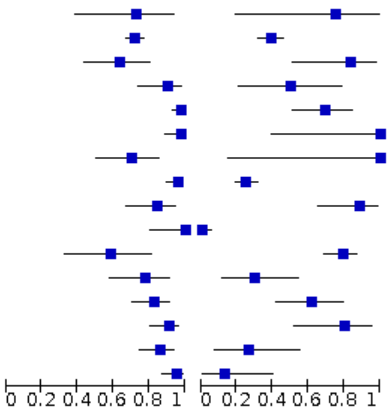
All

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Adil 2017	70	2	12	6	0.85 [0.76, 0.92]	0.75 [0.35, 0.97]
Adil 2017	69	1	13	7	0.84 [0.74, 0.91]	0.88 [0.47, 1.00]
Allegra 2009	26	1	4	14	0.87 [0.69, 0.96]	0.93 [0.68, 1.00]
Amirchaghmaghi 2018	21	29	0	4	1.00 [0.84, 1.00]	0.12 [0.03, 0.28]
Awan 2011	37	61	7	11	0.84 [0.70, 0.93]	0.15 [0.08, 0.26]
Awan 2012	34	52	10	20	0.77 [0.62, 0.89]	0.28 [0.18, 0.40]
Awan 2012	23	22	18	29	0.56 [0.40, 0.72]	0.57 [0.42, 0.71]
Baeten 2018	32	4	4	13	0.89 [0.74, 0.97]	0.76 [0.50, 0.93]
Cancela-Rodriguez 2011	19	35	10	96	0.66 [0.46, 0.82]	0.73 [0.65, 0.81]
Chainani-Wu 2015	2	13	15	40	0.12 [0.01, 0.36]	0.75 [0.62, 0.86]
Chainani-Wu 2015	2	7	15	46	0.12 [0.01, 0.36]	0.87 [0.75, 0.95]
Chaudhari 2013	72	1	4	5	0.95 [0.87, 0.99]	0.83 [0.36, 1.00]
Chaudhari 2013	67	2	9	4	0.88 [0.79, 0.94]	0.67 [0.22, 0.96]
Chen 2007	26	9	3	20	0.90 [0.73, 0.98]	0.69 [0.49, 0.85]
Cheng 2003a	16	2	9	3	0.64 [0.43, 0.82]	0.60 [0.15, 0.95]
Cheng 2003b	24	1	5	0	0.83 [0.64, 0.94]	0.00 [0.00, 0.97]
Chiang 2019	56	32	7	25	0.89 [0.78, 0.95]	0.44 [0.31, 0.58]
Delavarian 2010	16	0	2	8	0.89 [0.65, 0.99]	1.00 [0.63, 1.00]
Du 2007	31	25	2	70	0.94 [0.80, 0.99]	0.74 [0.64, 0.82]
Epstein 2008	30	9	12	28	0.71 [0.55, 0.84]	0.76 [0.59, 0.88]
Farah 2007	10	45	0	0	1.00 [0.69, 1.00]	0.00 [0.00, 0.08]
Farah 2012	11	29	15	63	0.42 [0.23, 0.63]	0.68 [0.58, 0.78]
Fontes 2013	157	0	1	6	0.99 [0.97, 1.00]	1.00 [0.54, 1.00]
Ganga 2017	19	59	6	116	0.76 [0.55, 0.91]	0.66 [0.59, 0.73]
Guneri 2011	14	22	1	6	0.93 [0.68, 1.00]	0.21 [0.08, 0.41]
Gupta 2007	44	3	2	47	0.96 [0.85, 0.99]	0.94 [0.83, 0.99]
Hanken 2013	47	8	1	4	0.98 [0.89, 1.00]	0.33 [0.10, 0.65]
Jayasinghe 2020	28	7	13	12	0.68 [0.52, 0.82]	0.63 [0.38, 0.84]
Johnson 2019	55	9	11	25	0.83 [0.72, 0.91]	0.74 [0.56, 0.87]
Kämmerer 2013	17	0	14	57	0.55 [0.36, 0.73]	1.00 [0.94, 1.00]
Kämmerer 2013	21	0	9	57	0.70 [0.51, 0.85]	1.00 [0.94, 1.00]
Kämmerer 2013	23	0	7	57	0.77 [0.58, 0.90]	1.00 [0.94, 1.00]
Kaur 2016	40	2	8	46	0.83 [0.70, 0.93]	0.96 [0.86, 0.99]
Kaur 2016	33	0	15	48	0.69 [0.54, 0.81]	1.00 [0.93, 1.00]
Koch 2011a	108	4	7	63	0.94 [0.88, 0.98]	0.94 [0.85, 0.98]
Koch 2011b	31	1	2	44	0.94 [0.80, 0.99]	0.98 [0.88, 1.00]
Leunig 2000	97	22	3	38	0.97 [0.91, 0.99]	0.63 [0.50, 0.75]
Liu 2019	68	22	18	95	0.79 [0.69, 0.87]	0.81 [0.73, 0.88]
Ma 2014	19	3	3	27	0.86 [0.65, 0.97]	0.90 [0.73, 0.98]
Mashberg 1980	101	8	16	110	0.86 [0.79, 0.92]	0.93 [0.87, 0.97]
McIntosh 2009	7	12	2	29	0.78 [0.40, 0.97]	0.71 [0.54, 0.84]
Mehrotra 2008	30	3	4	42	0.88 [0.73, 0.97]	0.93 [0.82, 0.99]
Mehrotra 2010	6	88	6	56	0.50 [0.21, 0.79]	0.39 [0.31, 0.47]
Mehrotra 2010	0	24	4	74	0.00 [0.00, 0.60]	0.76 [0.66, 0.84]
Mehrotra 2011	26	5	1	47	0.96 [0.81, 1.00]	0.90 [0.79, 0.97]
Mojsa 2012	7	25	0	9	1.00 [0.59, 1.00]	0.26 [0.13, 0.44]
Nagaraju 2010	55	4	0	1	1.00 [0.94, 1.00]	0.20 [0.01, 0.72]
Nanayakkara 2016	165	0	2	14	0.99 [0.96, 1.00]	1.00 [0.77, 1.00]
Nanayakkara 2016	154	0	13	14	0.92 [0.87, 0.96]	1.00 [0.77, 1.00]
Navone 2004	39	2	6	31	0.87 [0.73, 0.95]	0.94 [0.80, 0.99]
Navone 2008	71	12	2	73	0.97 [0.90, 1.00]	0.86 [0.77, 0.92]
Ng 2012	27	3	45	96	0.38 [0.26, 0.50]	0.97 [0.91, 0.99]
Onizawa 1999	75	2	5	42	0.94 [0.86, 0.98]	0.95 [0.85, 0.99]
Onofre 2001	10	13	3	24	0.77 [0.46, 0.95]	0.65 [0.47, 0.80]
Petruzzi 2014	24	10	6	16	0.80 [0.61, 0.92]	0.62 [0.41, 0.80]
Petruzzi 2014	21	11	9	15	0.70 [0.51, 0.85]	0.58 [0.37, 0.77]
Rahman 2012	19	13	8	46	0.70 [0.50, 0.86]	0.78 [0.65, 0.88]
Rahman 2012	18	11	9	48	0.67 [0.46, 0.83]	0.81 [0.69, 0.90]
Scheer 2011	12	10	0	42	1.00 [0.74, 1.00]	0.81 [0.67, 0.90]
Scheifele 2004	24	4	2	66	0.92 [0.75, 0.99]	0.94 [0.86, 0.98]
Sciubba 1999	102	14	0	182	1.00 [0.96, 1.00]	0.93 [0.88, 0.96]
Seijas-Naya 2012	8	1	3	12	0.73 [0.39, 0.94]	0.92 [0.64, 1.00]
Sharwani 2006a	28	8	3	30	0.90 [0.74, 0.98]	0.79 [0.63, 0.90]
Sharwani 2006b	8	1	3	3	0.73 [0.39, 0.94]	0.75 [0.19, 0.99]
Shi 2019	239	113	92	73	0.72 [0.67, 0.77]	0.39 [0.32, 0.47]



Test 1. (Continued)

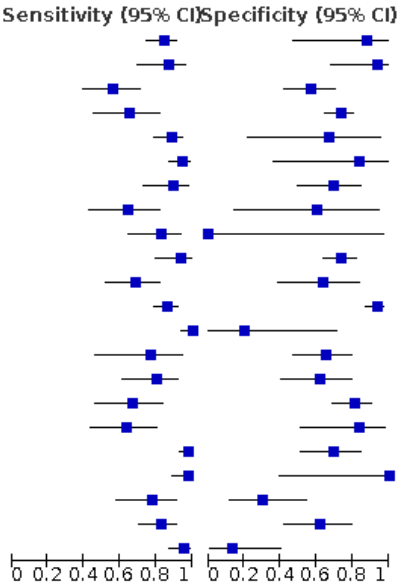
Sharwani 2006b	8	1	3	3	0.73 [0.39, 0.94]	0.75 [0.19, 0.99]
Shi 2019	239	113	92	73	0.72 [0.67, 0.77]	0.39 [0.32, 0.47]
Shukla 2018	19	2	11	10	0.63 [0.44, 0.80]	0.83 [0.52, 0.98]
Shukla 2018	27	6	3	6	0.90 [0.73, 0.98]	0.50 [0.21, 0.79]
Silverman 1984	97	10	2	23	0.98 [0.93, 1.00]	0.70 [0.51, 0.84]
Singh 2015	45	0	1	4	0.98 [0.88, 1.00]	1.00 [0.40, 1.00]
Skandarajah 2017	21	0	9	2	0.70 [0.51, 0.85]	1.00 [0.16, 1.00]
Svirsky 2002	93	150	4	51	0.96 [0.90, 0.99]	0.25 [0.20, 0.32]
Trakroo 2015	27	2	5	16	0.84 [0.67, 0.95]	0.89 [0.65, 0.99]
Ujaoney 2012	17	81	0	1	1.00 [0.80, 1.00]	0.01 [0.00, 0.07]
Ujaoney 2012	10	17	7	65	0.59 [0.33, 0.82]	0.79 [0.69, 0.87]
Upadhyay 2011	21	14	6	6	0.78 [0.58, 0.91]	0.30 [0.12, 0.54]
Warnakulasuriya 1996	47	11	10	18	0.82 [0.70, 0.91]	0.62 [0.42, 0.79]
Yamamoto 2017	58	3	6	12	0.91 [0.81, 0.96]	0.80 [0.52, 0.96]
Yamamoto 2017	55	11	9	4	0.86 [0.75, 0.93]	0.27 [0.08, 0.55]
Yamamoto 2017	61	13	3	2	0.95 [0.87, 0.99]	0.13 [0.02, 0.40]



Test 2. Vital staining with application method covariate

Vital staining with application method covariate

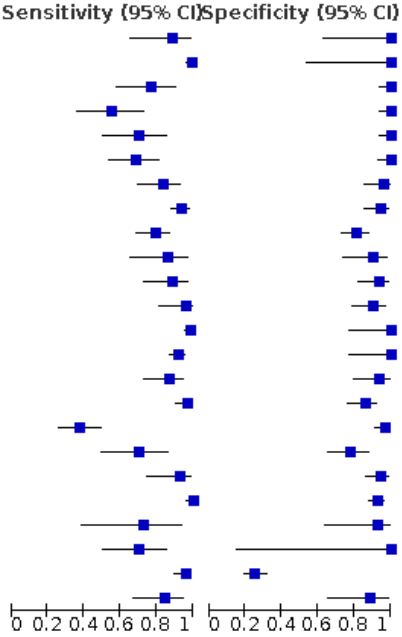
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Adil 2017	69	1	13	7	0.84 [0.74, 0.91]	0.88 [0.47, 1.00]
Allegra 2009	26	1	4	14	0.87 [0.69, 0.96]	0.93 [0.68, 1.00]
Awan 2012	23	22	18	29	0.56 [0.40, 0.72]	0.57 [0.42, 0.71]
Cancela-Rodriguez 2011	19	35	10	96	0.66 [0.46, 0.82]	0.73 [0.65, 0.81]
Chaudhari 2013	67	2	9	4	0.88 [0.79, 0.94]	0.67 [0.22, 0.96]
Chaudhari 2013	72	1	4	5	0.95 [0.87, 0.99]	0.83 [0.36, 1.00]
Chen 2007	26	9	3	20	0.90 [0.73, 0.98]	0.69 [0.49, 0.85]
Cheng 2003a	16	2	9	3	0.64 [0.43, 0.82]	0.60 [0.15, 0.95]
Cheng 2003b	24	1	5	0	0.83 [0.64, 0.94]	0.00 [0.00, 0.97]
Du 2007	31	25	2	70	0.94 [0.80, 0.99]	0.74 [0.64, 0.82]
Jayasinghe 2020	28	7	13	12	0.68 [0.52, 0.82]	0.63 [0.38, 0.84]
Mashberg 1980	101	8	16	110	0.86 [0.79, 0.92]	0.93 [0.87, 0.97]
Nagaraju 2010	55	4	0	1	1.00 [0.94, 1.00]	0.20 [0.01, 0.72]
Onofre 2001	10	13	3	24	0.77 [0.46, 0.95]	0.65 [0.47, 0.80]
Petruzzi 2014	24	10	6	16	0.80 [0.61, 0.92]	0.62 [0.41, 0.80]
Rahman 2012	18	11	9	48	0.67 [0.46, 0.83]	0.81 [0.69, 0.90]
Shukla 2018	19	2	11	10	0.63 [0.44, 0.80]	0.83 [0.52, 0.98]
Silverman 1984	97	10	2	23	0.98 [0.93, 1.00]	0.70 [0.51, 0.84]
Singh 2015	45	0	1	4	0.98 [0.88, 1.00]	1.00 [0.40, 1.00]
Upadhyay 2011	21	14	6	6	0.78 [0.58, 0.91]	0.30 [0.12, 0.54]
Warnakulasuriya 1996	47	11	10	18	0.82 [0.70, 0.91]	0.62 [0.42, 0.79]
Yamamoto 2017	61	13	3	2	0.95 [0.87, 0.99]	0.13 [0.02, 0.40]



Test 3. Cytology with covariate

Cytology with covariate

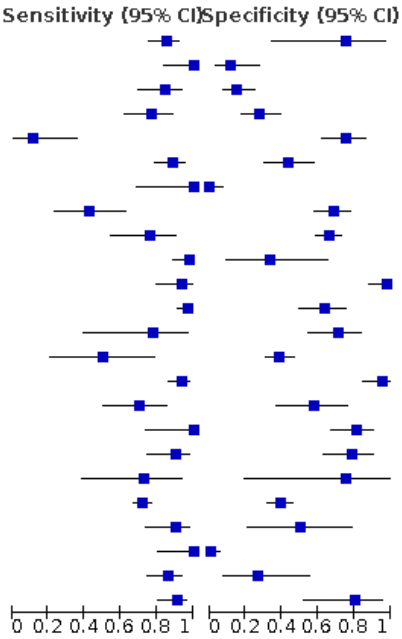
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Delavarian 2010	16	0	2	8	0.89 [0.65, 0.99]	1.00 [0.63, 1.00]
Fontes 2013	157	0	1	6	0.99 [0.97, 1.00]	1.00 [0.54, 1.00]
Kämmerer 2013	23	0	7	57	0.77 [0.58, 0.90]	1.00 [0.94, 1.00]
Kämmerer 2013	17	0	14	57	0.55 [0.36, 0.73]	1.00 [0.94, 1.00]
Kämmerer 2013	21	0	9	57	0.70 [0.51, 0.85]	1.00 [0.94, 1.00]
Kaur 2016	33	0	15	48	0.69 [0.54, 0.81]	1.00 [0.93, 1.00]
Kaur 2016	40	2	8	46	0.83 [0.70, 0.93]	0.96 [0.86, 0.99]
Koch 2011a	108	4	7	63	0.94 [0.88, 0.98]	0.94 [0.85, 0.98]
Liu 2019	68	22	18	95	0.79 [0.69, 0.87]	0.81 [0.73, 0.88]
Ma 2014	19	3	3	27	0.86 [0.65, 0.97]	0.90 [0.73, 0.98]
Mehrotra 2008	30	3	4	42	0.88 [0.73, 0.97]	0.93 [0.82, 0.99]
Mehrotra 2011	26	5	1	47	0.96 [0.81, 1.00]	0.90 [0.79, 0.97]
Nanayakkara 2016	165	0	2	14	0.99 [0.96, 1.00]	1.00 [0.77, 1.00]
Nanayakkara 2016	154	0	13	14	0.92 [0.87, 0.96]	1.00 [0.77, 1.00]
Navone 2004	39	2	6	31	0.87 [0.73, 0.95]	0.94 [0.80, 0.99]
Navone 2008	71	12	2	73	0.97 [0.90, 1.00]	0.86 [0.77, 0.92]
Ng 2012	27	3	45	96	0.38 [0.26, 0.50]	0.97 [0.91, 0.99]
Rahman 2012	19	13	8	46	0.70 [0.50, 0.86]	0.78 [0.65, 0.88]
Scheifele 2004	24	4	2	66	0.92 [0.75, 0.99]	0.94 [0.86, 0.98]
Sciubba 1999	102	14	0	182	1.00 [0.96, 1.00]	0.93 [0.88, 0.96]
Seijas-Naya 2012	8	1	3	12	0.73 [0.39, 0.94]	0.92 [0.64, 1.00]
Skandarajah 2017	21	0	9	2	0.70 [0.51, 0.85]	1.00 [0.16, 1.00]
Svirsky 2002	93	150	4	51	0.96 [0.90, 0.99]	0.25 [0.20, 0.32]
Trakroo 2015	27	2	5	16	0.84 [0.67, 0.95]	0.89 [0.65, 0.99]



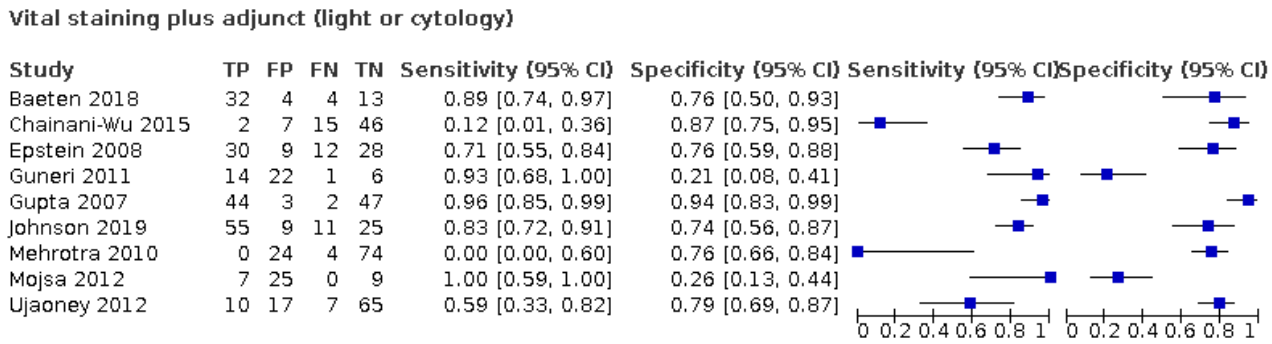
Test 4. Light-based with covariate

Light-based with covariate

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Adil 2017	70	2	12	6	0.85 [0.76, 0.92]	0.75 [0.35, 0.97]
Amirchaghmaghi 2018	21	29	0	4	1.00 [0.84, 1.00]	0.12 [0.03, 0.28]
Awan 2011	37	61	7	11	0.84 [0.70, 0.93]	0.15 [0.08, 0.26]
Awan 2012	34	52	10	20	0.77 [0.62, 0.89]	0.28 [0.18, 0.40]
Chainani-Wu 2015	2	13	15	40	0.12 [0.01, 0.36]	0.75 [0.62, 0.86]
Chiang 2019	56	32	7	25	0.89 [0.78, 0.95]	0.44 [0.31, 0.58]
Farah 2007	10	45	0	0	1.00 [0.69, 1.00]	0.00 [0.00, 0.08]
Farah 2012	11	29	15	63	0.42 [0.23, 0.63]	0.68 [0.58, 0.78]
Ganga 2017	19	59	6	116	0.76 [0.55, 0.91]	0.66 [0.59, 0.73]
Hanken 2013	47	8	1	4	0.98 [0.89, 1.00]	0.33 [0.10, 0.65]
Koch 2011b	31	1	2	44	0.94 [0.80, 0.99]	0.98 [0.88, 1.00]
Leunig 2000	97	22	3	38	0.97 [0.91, 0.99]	0.63 [0.50, 0.75]
McIntosh 2009	7	12	2	29	0.78 [0.40, 0.97]	0.71 [0.54, 0.84]
Mehrotra 2010	6	88	6	56	0.50 [0.21, 0.79]	0.39 [0.31, 0.47]
Onizawa 1999	75	2	5	42	0.94 [0.86, 0.98]	0.95 [0.85, 0.99]
Petruzzi 2014	21	11	9	15	0.70 [0.51, 0.85]	0.58 [0.37, 0.77]
Scheer 2011	12	10	0	42	1.00 [0.74, 1.00]	0.81 [0.67, 0.90]
Sharwani 2006a	28	8	3	30	0.90 [0.74, 0.98]	0.79 [0.63, 0.90]
Sharwani 2006b	8	1	3	3	0.73 [0.39, 0.94]	0.75 [0.19, 0.99]
Shi 2019	239	113	92	73	0.72 [0.67, 0.77]	0.39 [0.32, 0.47]
Shukla 2018	27	6	3	6	0.90 [0.73, 0.98]	0.50 [0.21, 0.79]
Ujaoney 2012	17	81	0	1	1.00 [0.80, 1.00]	0.01 [0.00, 0.07]
Yamamoto 2017	55	11	9	4	0.86 [0.75, 0.93]	0.27 [0.08, 0.55]
Yamamoto 2017	58	3	6	12	0.91 [0.81, 0.96]	0.80 [0.52, 0.96]



Test 6. Vital staining plus adjunct (light or cytology)



ADDITIONAL TABLES

Table 1. Index tests for oral cancer and OPMDs

Test	Characteristics	Classification of response	Other information
Vital staining (e.g. toluidine blue, tolonium chloride)	Vital staining refers to the use of dyes such as toluidine blue or tolonium chloride to stain oral mucosa tissues for OPMD or malignancy (Leston 2010; Lingen 2008; Patton 2008). The procedure is as follows: • pre-rinse with acetic acid • rinse with water • apply toluidine blue • post-rinse with acetic acid • rinse with water • observe mucosa to check for staining	The result of the test is classified as positive if tissue is stained and negative if no tissue is stained, or equivocal if no definitive result can be obtained	**Advantages:** ability to define areas that could be malignant or abnormal but cannot be seen; assess the extent of the OPMD for excision **Disadvantages:** benign inflammatory lesions are subject to stain; possibility of failure of some cancerous lesions to stain; possibility of failure of some dysplastic lesions (particularly those with a lower grade or with a thick keratotic surface) to stain; variation in test performance depending on how thorough the test procedures are followed; contraindicated in those who are known to be allergic to iodine
Brush cytology (e.g. OralCDx brush biopsy)	Brush cytology refers to the microscopic assessment and interpretation of cell samples from OPMD that are flaked off from the oral mucosa by the brushing, smearing, scraping or lavage to collect cell samples, which are then sealed on glass slides. They are then analysed using an imaging system that assesses the sampled cells (Leston 2010; Lingen 2008; Patton 2008)	Following analysis, cytopathologists classify test results as positive, atypical, or negative.	**Advantages:** include the ability to collect information from, and detect large or multiple lesions and to access "the basement membrane collecting cells from all three epithelial layers of the oral mucosa. The liquid-based cytology reduces the problems relating to sampling and fixation and presents a better cytological morphology" (Divani 2009) **Disadvantages:** smaller or less obvious lesions may be overlooked; difficulties in detecting lesions when there is necrosis or coagulated blood; inadequate training of operators (Divani 2009); cells are potentially seen out of context
Light-based detection (chemiluminescence e.g. ViziLite plus, tissue fluores-	Light-based systems to identify malignant and potentially malignant lesions and to highlight their presence through tissue reflectance (Leston 2010; Lingen	The result of the test is classed as negative if the appearance of the epithelium is lightly	**Advantages:** simple to use; non-invasive; do not require consumable reagents; provide real-time results; can be performed by a wide range of operators after a short training period

Table 1. Index tests for oral cancer and OPMDs *(Continued)*

cence imaging e.g. ViziLite, Microlux DL; VELscope, Identafi 3000; tissue fluorescence spectroscopy)	2008; Patton 2008) e.g. using Microlux DL, the procedure is as follows (Lingen 2008): • pre-rinse with acetic acid • use blue-light source to visually assess the oral cavity ViziLite Plus also provides a tolonium chloride solution (toluidine blue) to aid in the marking of the lesion for biopsy once the light source is removed	bluish white and positive if the appearance of the epithelium is distinctly white (acetowhite)	**Disadvantages:** the necessity of a dark environment; high initial set up (for VELscope) or recurrent costs (for ViziLite in low-income countries); lack of permanent record unless photographed; inability to objectively measure visualisation results
Blood and saliva analysis	These novel technologies are at an early stage of development and evaluation. Analysis of blood or saliva samples which tests for the presence of biomarkers of OPMD and oral cancer (Brinkmann 2011; Lee 2009; Li 2006)	Cut-off probabilities vary widely and are dependent on the individual biomarker or combination of biomarkers examined	**Advantages:** non-invasive (saliva tests) or minimally invasive (blood tests) **Disadvantages:** there is a tendency for the estimated diagnostic accuracy of new health technologies to decline over time as evidence from independent evaluations accumulate (Wyatt 1995). This bias, which can be substantial, has been demonstrated in other domains, e.g. acute abdominal pain (Liu 2006) and clinical decision support systems (Garg 2005). Promising biomarker tests in several clinical areas were eventually been shown to be disappointing (Buchen 2011). It remains to be seen whether this is the case with oral cancer and OPMDs

OPMD = oral potentially malignant disorders.

Table 2. Indicators for the assessment of quality (QUADAS-2)

Domain	Patient selection	Index test	Reference standard	Flow and timing
Description	Describe methods of patient selection. Describe included patients (characteristics, prior testing, presentation and severity of the target condition (class), intended use of index test and setting)	Describe the index test(s) and how it was conducted and interpreted. Describe the sequence of tests, any training or calibration of clinicians (levels of agreement should be reported; where this is measured by the kappa statistic, acceptable values range from 0.61 (moderate agreement) to 1.00 (almost perfect agreement) (Landis 1977)), any procedures taken to ensure blinding of examiners, post-hoc or a priori threshold specification, any conflict of interest or commercial funding. Methods of site selection should be clearly documented	Describe the reference standard and how it was conducted and interpreted. Ideally, the biopsied tissue should be examined by more than 1 pathologist. If there is a lack of agreement any methods for reaching consensus should be clearly documented. Any measures taken to ensure pathologists were blinded to the results of the index tests should be documented, along with the sequence of reference and index tests. Methods of site selection should be clearly documented	Describe the characteristics and proportion of patients who did not receive the index test(s) and/or reference standard, who received a reference standard other than the biopsy, or who were excluded from the 2 x 2 table (refer to flow diagram). Describe the time interval and any interventions between index test(s) and reference standard. The length of time between the index test and reference standard should be short in most cases. If the period elapsed between index test and reference

Table 2. Indicators for the assessment of quality (QUADAS-2) *(Continued)*

				standard is greater than 2 weeks then this will be considered an unacceptable delay
Signalling questions (Yes/No/Unclear)	Was a consecutive or random sample of patients enrolled? • Classify as **'Yes'** if consecutive patients or a random sample of individuals were recruited • Classify as **'No'** if non-consecutive patients or a non-random sample of individuals were recruited • Classify as **'Unclear'** if patient selection was not clearly described	Was calibration of examiners undertaken and results reported? • Classify as **'Yes'** if the examiners participated in dedicated training and calibration was reported to an acceptable standard • Classify as **'No'** if the examiners did not participate in dedicated training or was not assessed, or training was undertaken but calibration was not to an acceptable standard • Classify as **'Unclear'** if the information on training and calibration was not stated	Is the reference standard likely to correctly classify the target condition? • Classify as **'Yes'** if the biopsy was independently confirmed by at least 2 qualified pathologists • Classify as **'No'** if the biopsy was not independently confirmed by at least 2 qualified pathologists, or there was lack of agreement between pathologists • Classify as **'Unclear'** if the study does not state who confirmed the biopsy	Was there an appropriate time interval between the index test(s) and reference standard? • Classify as **'Yes'** if the delay between the index test(s) and reference standard is considered acceptable for most participants • Classify as **'No'** if the delay between the index test(s) and reference standard is considered unacceptable for most participants • Classify as **'Unclear'** if the delay between the index test(s) and reference standard is not explicitly stated
-	Did the study avoid inappropriate exclusions? • Classify as **'Yes'** if patients with either class I or class II lesions were recruited • Classify as **'No'** if only patients with class I lesions were recruited • Classify as **'Unclear'** if class of lesions was not clearly described	Were the index test results interpreted without knowledge of the results of the reference standard? • Classify as **'Yes'** if interpreters of index test results clearly do not know results of biopsy/histopathology • Classify as **'No'** if interpreters of index test results clearly know results of biopsy/histopathology • Classify as **'Unclear'** if study did not provide any information on whether interpreters of index tests were blinded to biopsy/histopathology	Were the reference standard results interpreted without knowledge of the results of the index test? • Classify as **'Yes'** if pathologists clearly do not know the index test results when interpreting biopsied tissues • Classify as **'No'** if pathologists know the results of index test results when interpreting biopsied tissues • Classify as **'Unclear'** if the study did not provide any information on whether the pathologists were blinded to the index test results	Did all patients receive the same reference standard? • Classify as **'Yes'** if the same reference standard was used in all participants • Classify as **'No'** if the same reference standard was not used in all participants • Classify as **'Unclear'** if it is unclear whether different reference standards were used

Table 2. Indicators for the assessment of quality (QUADAS-2) *(Continued)*

-	-	Where multiple index tests were used, were the results of the second index test interpreted without knowledge of the results of the first index test? • Classify as **'Yes'** if index test results were interpreted without knowledge • Classify as **'No'** if the index test results were interpreted with knowledge • Classify as **'Unclear'** if it is unclear whether the results of the second index test were interpreted without knowledge of the results of the first index test?	-	Were all patients included in the analysis? • Classify as **'Yes'** if all patients were included in the analysis • Classify as **'No'** is only some patients were included in the analysis • Classify as **'Unclear'** if it is unclear whether all patients were included in the analysis
-	-	If a threshold was used, was it pre-specified? • Classify as **'Yes'** if the threshold was pre-specified • Classify as **'No'** if the threshold was not pre-specified • Classify as **'Unclear'** if it is unclear whether the threshold was pre-specified	-	-
-	-	Were any conflicts of interest stated? • Classify as **'Yes'** if the study declared no conflict of interest • Classify as **'No'** if the study if the study declared a conflict of interest • Classify as **'Unclear'** there was no information on conflict of interest	-	-
Risk of bias: High/Low/Unclear	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?
Concerns regarding applicability: High/Low/Unclear	Are there concerns that the included patients do not match the review question?	Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Are there concerns that the target condition as defined by the reference standard does not match the review question?	-

QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies 2.

APPENDICES

Appendix 1. MEDLINE Ovid search strategy

1. exp Mouth/
2. Cheek/

3. or/1-2
4. exp Carcinoma, squamous cell/di
5. exp Precancerous conditions/di
6. (tumor$ or tumour$ or cancer$ or carcinoma$ or carcinogen$ or neoplas$ or malignan$ or metasta$ or dysplas$ or lesion$ or ulcer $).tw,ot.
7. (pre-cancer$ or precancer$ or premalignan$ or precursor$ or "lichen planus" or leukoplakia or "submucous fibrosis" or (actinic adj2 keratosis) or candidiasis or erythroplakia or erythroplas$ or erythroleukoplakia or hyperplas$ or hyperkeratos$).tw,ot.
8. or/4-7
9. 3 and 8
10. exp Mouth neoplasms/di
11. Lichen Planus, Oral/di
12. Oral submucous fibrosis/di
13. Oral candidiasis/di
14. ((oral$ or mouth$ or bucca$ or "oral cavit$" or (oral adj mucosa$) or (mouth adj mucosa$) or lip or lips or tongue$ or gingiv$ or palat$ or cheek$ or "intra oral$" or intraoral$ or gum or gums or labial$) adj3 (tumor$ or tumour$ or cancer$ or carcinoma$ or carcinogen$ or neoplas $ or malignan$ or metasta$ or dysplas$ or lesion$ or ulcer$ or pre-cancer$ or precancer$ or premalignan$ or precursor$ or "lichen planus" or leukoplakia or "submucous fibrosis" or (actinic adj2 keratosis) or candidiasis or erythroplakia or erythroplas$ or erythroleukoplakia or hyperplas$ or hyperkerato$)).tw,ot.
15. or/10-14
16. 9 or 15
17. Cytodiagnosis/
18. Cytological techniques/
19. Cytophotometry/
20. (brush adj3 biops$).tw,ot.
21. ("oral cdx" or oralcdx).tw,ot.
22. ("modified liquid based cytology" or (exfoliat$ adj3 cytolog$)).tw,ot.
23. (brush$ and (cytodiagnosis or cytopathology)).tw,ot.
24. Tolonium chloride/
25. Coloring agents/du
26. ("tolonium chloride" or "tolu?dine blue" or "tolu?dine b" or tblue or t-blue).tw,ot.
27. (tolu?dine adj6 (dye$ or rins$ or stain$ or wash$)).tw,ot.
28. exp Luminescence/
29. Fluorescence/
30. Spectrometry, fluorescence/
31. exp Luminescent Agents/
32. Light/
33. Tomography, Optical Coherence/
34. (visual$ adj5 ("light emitting diode" or "blue spectrum" or LED or luminous$)).tw,ot.
35. (visuali?ation adj3 adjunct$).tw,ot.
36. (vizilite or microlux$ or orascoptic or velscope).tw,ot.
37. lumenoscop$.tw,ot.
38. ((tumor$ or tumour$ or cancer$ or carcinoma$ or neoplas$ or carcinogen$ or malignan$ or metata$ or lesion$ or ulcer$) adj5 (fluorescen$ or autofluorescen$ or luminescen$ or chemiluminescen$)).tw,ot.
39. (tissue adj3 reflect$).tw,ot.
40. Spectrophotometry/
41. Acetic acid/
42. (acetic acid adj3 (wash$ or rins$)).tw,ot.
43. acetowhite.tw,ot.
44. Saliva/an, ch
45. Tumor Markers, Biological/an
46. (("tumo?r marker$" or "neoplas$ marker$") adj3 (blood or saliva)).tw,ot.
47. ((analy$ or screen$ or test$ or examin$) adj3 (blood or saliva)).tw,ot.
48. Diagnosis, Oral/
49. Mass screening/
51. screen$.mp.
50. Physical examination/
51. ((oral$ or mouth$) adj5 (exam$ or histolog$ or check$ or inspect$)).tw,ot.
52. (visual$ adj3 (exam$ or inspect$ or screen$)).tw,ot.
53. or/17-52
54. 16 and 53

Appendix 2. Embase Ovid search strategy

1. exp Mouth/
2. Cheek/
3. or/1-2
4. exp Squamous cell carcinoma/di
5. exp Precancer/di
6. (tumor$ or tumour$ or cancer$ or carcinoma$ or carcinogen$ or neoplas$ or malignan$ or metasta$ or dysplas$ or lesion$ or ulcer $).tw,ot.
7. (pre-cancer$ or precancer$ or premalignan$ or precursor$ or "lichen planus" or leukoplakia or "submucous fibrosis" or (actinic adj2 keratosis) or candidiasis or erythroplakia or erythroplas$ or erythroleukoplakia or hyperplas$ or hyperkeratos$).tw,ot.
8. or/4-7
9. 3 and 8
10. exp Mouth tumor/di
11. Lichen planus/di
12. Thrush/di
13. ((oral$ or mouth$ or bucca$ or "oral cavit$" or (oral adj mucosa$) or (mouth adj mucosa$) or lip or lips or tongue$ or gingiv$ or palat$ or cheek$ or "intra oral$" or intraoral$ or gum or gums or labial$) adj3 (tumor$ or tumour$ or cancer$ or carcinoma$ or carcinogen$ or neoplas $ or malignan$ or metasta$ or dysplas$ or lesion$ or ulcer$ or pre-cancer$ or precancer$ or premalignan$ or precursor$ or "lichen planus" or leukoplakia or "submucous fibrosis" or (actinic adj2 keratosis) or candidiasis or erythroplakia or erythroplas$ or erythroleukoplakia or hyperplas$ or hyperkerato$)).tw,ot.
14. or/10-13
15. 9 or 14
16. Cancer cytodiagnosis/
17. Cytophotometry/
18. (brush adj3 biops$).tw,ot.
19. ("oral cdx" or oralcdx).tw,ot.
20. ("modified liquid based cytology" or (exfoliat$ adj3 cytolog$)).tw,ot.
21. (brush$ and (cytodiagnosis or cytopathology)).tw,ot.
22. Tolonium chloride/
23. Coloring agent/
24. ("tolonium chloride" or "tolu?dine blue" or "tolu?dine b" or tblue or t-blue).tw,ot.
25. (tolu?dine adj6 (dye$ or rins$ or stain$ or wash$)).tw,ot.
26. exp Luminescence/
27. Fluorescence/
28. Spectrofluorometry/
29. exp Luminescent Agents/
30. Light/
31. Tomography, Optical Coherence/
32. (visual$ adj5 ("light emitting diode" or "blue spectrum" or LED or luminous$)).tw,ot.
33. (visuali?ation adj3 adjunct$).tw,ot.
34. (vizilite or microlux$ or orascoptic or velscope).tw,ot.
35. lumenoscop$.tw,ot.
36. ((tumor$ or tumour$ or cancer$ or carcinoma$ or neoplas$ or carcinogen$ or malignan$ or metata$ or lesion$ or ulcer$) adj5 (fluorescen$ or autofluorescen$ or luminescen$ or chemiluminescen$)).tw,ot.
37. (tissue adj3 reflect$).tw,ot.
38. Spectrophotometry/
39. Acetic acid/
40. (acetic acid adj3 (wash$ or rins$)).tw,ot.
41. acetowhite.tw,ot.
42. Tumor Marker/
43. (("tumo?r marker$" or "neoplas$ marker$") adj3 (blood or saliva)).tw,ot.
44 ((analy$ or screen$ or test$ or examin$) adj3 (blood or saliva)).tw,ot.
45. Mass screening/
46. screen$.mp.
47. Physical examination/
48. ((oral$ or mouth$) adj5 (diagnos$ or exam$ or histolog$ or check$ or inspect$)).tw,ot.
49. (visual$ adj3 (exam$ or inspect$ or screen$)).tw,ot.
50. or/16-49
51. 15 and 50

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

Condition: oral cancer

Other terms: diagnosis or diagnose

Condition: oral cancer

Other terms: screen or screening

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

oral cancer AND diagnosis OR oral cancer AND diagnose OR oral cancer AND diagnostic

oral cancer AND screen OR oral cancer AND screening

Appendix 5. Search strategies used in the previous version of this review (April 2013)

Cochrane Oral Health's Trials Register search strategy

An updated search of the Cochrane Oral Health's Trials Register was conducted 30 April 2013 using the Cochrane Register of Studies software and the search strategy below:
#1 ((oral* or mouth* or bucca* or "oral cavit*" or "oral mucosa" or "mouth mucosa" or lip or lips or tongue* or gingiva* or palat* or cheek* or intra-oral* or intraoral* or gum or gums or labial*):ti,ab) AND (INREGISTER)
#2 ((tumour* or tumor* or cancer* or carcinoma* or carcinogen* or neoplas* or malignan* or metasta* or dysplas* or lesion* or ulcer* or precancer* or pre-cancer* or premalignan* or precursor* or "lichen planus" or leukoplakia or "submucous fibrosis" or "actinic keratosis" or candidiasis or erythroplakia or erythroplas* or erythroleukoplakia or hyperplas* or hyperkerato*):ti,ab) AND (INREGISTER)
#3 ((cytodiagnosis or cytophotometry or "brush biops*" or "oral cdx" or oralcdx or "modified liquid based cytology" or "exfoliat* cytolog*" or "tolonium chloride" or "toludine b*" or "toluidine b*" or tblue or t-blue or "toludine dye*" or "toludine rins*" or "toludine stain*" or "toludine wash*" or "toluidine dye*" or "toluidine rins*" or "toluidine stain*" or "toluidine wash*" or luminescence or fluorescen* or "light emitting diode*"):ti,ab) AND (INREGISTER)
#4 (((blood or saliva) AND (analys* or inspect* or test or examin*)):ti,ab) AND (INREGISTER)
#5 (("blue spectrum" or LED or luminous or "visual* adjunct*" or vizilite or microlux* or orascoptic or velscope or lumenoscope* or autofluorescen* or chemilumiescen* or spectrophotometr* or "acetic acid" or acetowhite or "tumor marker*" or "tumour marker*" or "neoplas* marker*"):ti,ab) AND (INREGISTER)
#6 ((diagnos* AND (exam* or histolog* or check* or screen*)):ti,ab) AND (INREGISTER)
#7 (#1 and #2) AND (INREGISTER)
#8 (#3 or #4 or #5 or #6) AND (INREGISTER)
#9 (#7 and #8) AND (INREGISTER)

A previous search was conducted in June 2011 using the Procite software and the search strategies below:
((oral* or mouth* or bucca* or "oral cavit*" or "oral mucosa" or "mouth mucosa" or lip or lips or tongue* or gingiva* or palat* or cheek* or intra-oral* or intraoral* or gum or gums or labial*) AND (tumour* or tumor* or cancer* or carcinoma* or carcinogen* or neoplas* or malignan* or metasta* or dysplas* or lesion* or ulcer* or precancer* or pre-cancer* or premalignan* or precursor* or "lichen planus" or leukoplakia or "submucous fibrosis" or "actinic keratosis" or candidiasis or erythroplakia or erythroplas* or erythroleukoplakia or hyperplas* or hyperkerato*) AND (cytodiagnosis or cytophotometry or "brush biops*" or "oral cdx" or oralcdx or "modified liquid based cytology" or "exfoliat* cytolog*" or "tolonium chloride" or "toludine b*" or "toluidine b*" or tblue or t-blue or "toludine dye*" or "toludine rins*" or "toludine stain*" or "toludine wash*" or "toluidine dye*" or "toluidine rins*" or "toluidine stain*" or "toluidine wash*" or luminescence or fluorescen* or "light emitting diode*"))((oral* or mouth* or bucca* or "oral cavit*" or "oral mucosa" or "mouth mucosa" or lip or lips or tongue* or gingiva* or palat* or cheek* or intra-oral* or intraoral* or gum or gums or labial*) AND (tumour* or tumor* or cancer* or carcinoma* or carcinogen* or neoplas* or malignan* or metasta* or dysplas* or lesion* or ulcer* or precancer* or pre-cancer* or premalignan* or precursor* or "lichen planus" or leukoplakia or "submucous fibrosis" or "actinic keratosis" or candidiasis or erythroplakia or erythroplas* or erythroleukoplakia or hyperplas* or hyperkerato*) AND ("blue spectrum" or LED or luminous or "visual* adjunct*" or vizilite or microlux* or orascoptic or velscope or lumenoscope* or autofluorescen* or chemilumiescen* or spectrophotometr* or "acetic acid" or acetowhite or "tumor marker*" or "tumour marker*" or "neoplas* marker*"))((oral* or mouth* or bucca* or "oral cavit*" or "oral mucosa" or "mouth mucosa" or lip or lips or tongue* or gingiva* or palat* or cheek* or intra-oral* or intraoral* or gum or gums or labial*) AND (tumour* or tumor* or cancer* or carcinoma* or carcinogen* or neoplas* or malignan* or metasta* or dysplas* or lesion* or ulcer* or precancer* or pre-cancer* or premalignan* or precursor* or "lichen planus" or leukoplakia or "submucous fibrosis" or "actinic keratosis" or candidiasis or erythroplakia or erythroplas* or erythroleukoplakia or hyperplas* or hyperkerato*) AND (diagno* and (blood or saliva) and (analys* or inspect* or test* or examin*)))((oral* or mouth* or bucca* or "oral cavit*" or "oral mucosa" or "mouth mucosa" or lip or lips or tongue* or gingiva* or palat* or cheek* or intra-oral* or intraoral* or gum or gums or labial*) AND (tumour* or tumor* or cancer* or carcinoma* or carcinogen* or neoplas* or malignan* or metasta* or dysplas* or lesion* or ulcer* or precancer* or pre-cancer* or premalignan* or precursor* or "lichen planus" or leukoplakia or "submucous fibrosis" or "actinic keratosis" or candidiasis or erythroplakia or erythroplas* or erythroleukoplakia or hyperplas* or hyperkerato*) AND (diagnos* AND (exam* or histolog* or check or inspect* or screen*)))

Cochrane Diagnostic Test Accuracy Register search strategy

((oral* or mouth* or bucca* or "oral cavit*" or "oral mucosa" or "mouth mucosa" or lip or lips or tongue* or gingiva* or palat* or cheek* or intra-oral* or intraoral* or gum or gums or labial*) AND (tumour* or tumor* or cancer* or carcinoma* or carcinogen* or neoplas* or malignan* or metasta* or dysplas* or lesion* or ulcer* or precancer* or pre-cancer* or premalignan* or precursor* or "lichen planus" or leukoplakia or "submucous fibrosis" or "actinic keratosis" or candidiasis or erythroplakia or erythroplas* or erythroleukoplakia or hyperplas* or hyperkerato*)

MEDION search strategy

Searched using the code C (malignancies), and screened the results for oral cancer terms.

WHAT'S NEW

Date	Event	Description
16 December 2021	Amended	Additional external source of support added

HISTORY

Protocol first published: Issue 12, 2012
Review first published: Issue 5, 2015

Date	Event	Description
29 March 2021	New search has been performed	Searches updated to 20 October 2020
29 March 2021	New citation required but conclusions have not changed	Review update including 23 new studies bringing the total to 63 included studies. Changes to author byline. Conclusions remain unchanged

CONTRIBUTIONS OF AUTHORS

All review authors collaborated in the conception and design of the review.
Developing the search strategy: Tanya Walsh (TW).
Selecting studies for inclusion: Richard Macey (RM), TW.
Extracting data: RM, TW.
Carrying out analysis: TW, RM.
Interpreting the analysis: TW, RM, Alexander R Kerr (ARK), Graham Ogden (GO), Mark Lingen (ML), Saman Warnakulasuriya (SW).
Drafting the final review: RM, TW, ARK, GO, ML, SW.
The final review was read and approved by all review authors.

DECLARATIONS OF INTEREST

Tanya Walsh: none known. I am Statistical Editor for Cochrane Oral Health.
Richard Macey: none known.
Alexander R Kerr: none known.
Mark Lingen: none known.
Graham Ogden: none known.
Saman Warnakulasuriya: none known. I was a co-author on three of the included studies but was not involved in selecting or assessing them.

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
- Scottish Dental Clinical Effectiveness Programme (SDCEP), UK

This review has been supported by SDCEP.

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.

- National Institute for Health Research (NIHR), UK

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

- Cochrane Oral Health Global Alliance, Other

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

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

- The discussion of medical imaging techniques as an alternative test has been removed from the Background.
- We included studies reporting at the lesion level.

INDEX TERMS

Medical Subject Headings (MeSH)

Bias; Biomarkers, Tumor [analysis] [blood]; Carcinoma, Squamous Cell [*diagnosis] [pathology]; Coloring Agents; Early Detection of Cancer; Light; Lip Neoplasms [diagnosis] [pathology]; Mouth [pathology]; Mouth Neoplasms [*diagnosis] [pathology]; Saliva [chemistry]; Sensitivity and Specificity

MeSH check words

Humans