



University of Dundee

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

Macey, Richard; Walsh, Tanya; Brocklehurst, Paul; Kerr, Alexander R.; Liu, Joseph L. Y.; Lingen, Mark W.

Published in:
Cochrane Database of Systematic Reviews

DOI:
[10.1002/14651858.CD010276.pub2](https://doi.org/10.1002/14651858.CD010276.pub2)

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

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

Citation for published version (APA):
Macey, R., Walsh, T., Brocklehurst, P., Kerr, A. R., Liu, J. L. Y., Lingen, M. W., Ogden, G. R., Warnakulasuriya, S., & Scully, C. (2015). Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions. *Cochrane Database of Systematic Reviews*, 2015(5), Article CD010276. <https://doi.org/10.1002/14651858.CD010276.pub2>

General rights

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

Take down policy

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



Cochrane
Library

Cochrane Database of Systematic Reviews

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

Macey R, Walsh T, Brocklehurst P, Kerr AR, Liu JLY, Lingen MW, Ogden GR, Warnakulasuriya S, Scully C

Macey R, Walsh T, Brocklehurst P, Kerr AR, Liu JLY, Lingen MW, Ogden GR, Warnakulasuriya S, Scully C.
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.

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	3
OBJECTIVES	5
METHODS	5
RESULTS	7
Figure 1.	8
Figure 2.	9
Figure 3.	10
Figure 4.	12
Figure 5.	13
Figure 6.	14
Figure 7.	15
Figure 8.	16
Figure 9.	17
DISCUSSION	20
AUTHORS' CONCLUSIONS	21
ACKNOWLEDGEMENTS	22
REFERENCES	22
CHARACTERISTICS OF STUDIES	28
DATA	140
Test 1. Vital staining.	140
Test 2. Cytology.	141
Test 3. Light-based.	142
Test 4. Vital staining plus adjunct (Light).	142
Test 5. Vital staining plus adjunct (Cytology).	143
ADDITIONAL TABLES	143
APPENDICES	148
CONTRIBUTIONS OF AUTHORS	152
DECLARATIONS OF INTEREST	152
SOURCES OF SUPPORT	153
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	153
INDEX TERMS	153

[Diagnostic Test Accuracy Review]

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

Richard Macey¹, Tanya Walsh¹, Paul Brocklehurst², Alexander R Kerr³, Joseph LY Liu⁴, Mark W Lingen⁵, Graham R Ogden⁶, Saman Warnakulasuriya⁷, Crispian Scully⁸

¹School of Dentistry, The University of Manchester, Manchester, UK. ²NWORTH CTU, Bangor University, Bangor, UK. ³Department of Oral and Maxillofacial Pathology, Radiology and Medicine, New York University College of Dentistry, New York, USA. ⁴University of Dundee, Dental Health Services Research Unit, Scottish Dental Clinical Effectiveness Programme, NHS Education for Scotland, Dundee, UK. ⁵Pritzker School of Medicine, Division of Biological Sciences, Department of Pathology, University of Chicago, Chicago, Illinois, USA. ⁶Division of Oral and Maxillofacial Clinical Sciences, School of Dentistry, University of Dundee, Dundee, UK. ⁷Clinical and Diagnostic Sciences, King's College London, London, UK. ⁸University College London, London, UK

Contact address: Richard Macey, School of Dentistry, The University of Manchester, Coupland 3 Building, Oxford Road, Manchester, M13 9PL, UK. richard.macey@manchester.ac.uk.

Editorial group: Cochrane Oral Health Group.

Publication status and date: New, published in Issue 5, 2015.

Review content assessed as up-to-date: 30 April 2013.

Citation: Macey R, Walsh T, Brocklehurst P, Kerr AR, Liu JLY, Lingen MW, Ogden GR, Warnakulasuriya S, Scully C. 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.

Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Oral squamous cell carcinoma is the most common form of malignancy of the lip and oral cavity, often being preceded by potentially malignant disorders (PMD). Early detection can reduce the malignant transformation of PMD and can improve the survival rate for oral cancer. The current standard of scalpel biopsy with histology is painful for patients and involves a delay whilst histology is completed; other tests are available that are unobtrusive and provide immediate results.

Objectives

Primary objective: To estimate the diagnostic accuracy of index tests for the detection of oral cancer and PMD of the lip and oral cavity, in people presenting with clinically evident lesions.

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

Search methods

The electronic databases were searched on 30 April 2013. We searched MEDLINE (OVID) (1946 to April 2013) and four other electronic databases (the Cochrane Diagnostic Test Accuracy Studies Register, the Cochrane Oral Health Group's Trials Register, EMBASE (OVID) and MEDION (Ovid)). There were no restrictions on language in the searches of 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 PMD or oral squamous cell carcinoma of the lip or oral cavity: vital staining, oral cytology, light-based detection and oral spectroscopy, blood or saliva analysis (which test for the presence of biomarkers in blood or saliva).

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

Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

1

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 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

We included 41 studies, recruiting 4002 participants, in this review. These studies evaluated the diagnostic accuracy of conventional oral examination with: vital staining (14 studies), oral cytology (13 studies), light-based detection or oral spectroscopy (13 studies). Six studies assessed two combined index tests. There were no eligible diagnostic accuracy studies evaluating blood or salivary sample analysis.

The summary estimates for vital staining obtained from the meta-analysis were sensitivity of 0.84 (95% CI 0.74 to 0.90) with specificity of 0.70 (0.59 to 0.79), with 14 studies were included in the meta-analysis. For cytology, sensitivity was 0.91 (0.81 to 0.96) and specificity was 0.91 (0.81 to 0.95) with 12 studies included in the meta-analysis. For light-based detection, sensitivity was 0.91 (0.77 to 0.97) and specificity was 0.58 (0.22 to 0.87) with 11 studies included in the meta-analysis. The relative test accuracy was assessed by adding covariates to the bivariate analysis, no difference in model fit was observed.

Authors' conclusions

The overall quality of the included studies was poor. None of the adjunctive tests can be recommended as a replacement for the currently used standard of a scalpel biopsy and histological assessment. Given the relatively high values of the summary estimates of sensitivity and specificity for cytology, this would appear to offer the most potential. Combined adjunctive tests involving cytology warrant further investigation.

PLAIN LANGUAGE SUMMARY

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

Review question

The current method of diagnosing cancer of the mouth or lips involves the surgical removal of a piece of affected tissue that is sent to a laboratory for histological examination using a microscope (scalpel biopsy). This is painful for patients and involves a delay. The aim of this review was to find out the accuracy of three alternative diagnostic tests that are less invasive and provide more timely results.

Background

Oral cancer (OSCC - oral squamous cell carcinoma) often occurs after a condition called PMD (potentially malignant disorder), which can sometimes progress to cancer. If conditions such as oral cancer or PMD are identified early enough, outcomes for patients can be improved.

Study characteristics

The search on which the evidence is based was carried out on 30 April 2013. Forty-one studies involving 4002 participants, published between 1980 and 2012, were included. Each participant underwent one of three diagnostic tests for oral cancer and PMD as well as the standard method of diagnosis by surgical biopsy. By comparing results, the researchers were able to evaluate the accuracy of each test as compared to surgical biopsy.

Three tests (index tests) were evaluated.

1. Vital stain - a liquid that can be used as a mouthrinse or applied straight on to a suspected area of the mouth. It is thought that any area that is coloured blue after rinsing with water has a high chance of being oral cancer or PMD.
2. Oral cytology - instead of using a scalpel to cut away a piece of the suspected area, a brush is used to remove cells that are sent to a laboratory for examination under a microscope.
3. Light-based detection - a special light shone in the mouth that is believed to make cancerous areas appear different to healthy areas.

There were no eligible studies that looked at the accuracy of tests of blood or saliva.

Key Results

The proportion of people with OSCC or PMD identified through surgical biopsy varied in the included studies. We used the middle value for the included studies to illustrate the implications of the different tests. A false negative result means that people that truly have oral cancer or PMD will be diagnosed as free from the condition, possibly to be correctly diagnosed at a later date when the condition will be more difficult to treat successfully. A false positive result would mean that people who did not truly have PMD or oral cancer would be diagnosed as having it and therefore unnecessarily undergo invasive treatment.

If vital staining was used to identify OSCC or PMD in a sample of 1000 people (of whom 500 truly have OSCC or PMD), then the condition would go undetected in 80 people (false negative result) and 150 people without the condition would be told they have the condition (false positive result). If cytology was used, the condition would go undetected in 45 people and 45 people without the condition would be told they have the condition. If a light-based detection method was used, the condition would go undetected in 45 people and 210 people without the condition would be told they have the condition.

Therefore, in terms of correctly classifying people, cytology was the most accurate of the three tests.

Quality of the evidence

There were shortcomings in many of the studies that put them at high risk of bias and so the key results should be interpreted with caution. The main concern was the ways in which people were selected to take part in the studies. When patients at particularly high or low risk of oral cancer are selected to participate then this can influence the results of the study. Additionally, there were studies where the results from the standard method of diagnosis ('index test') were not reported and the reasons for this were not explained.

Conclusion

None of the tests evaluated in this review that were additional to a visual examination can be recommended as a replacement for the currently used standard of a scalpel biopsy and histological assessment.

BACKGROUND

Target condition being diagnosed

The target conditions of interest are oral squamous cell carcinoma (OSCC) and potentially malignant disorders (PMD) of the lip and oral cavity. OSCC is the most common form of oral cavity cancer (Scully 2000a) and many are preceded by PMD. PMD represent a heterogeneous group of conditions including erythroplakia, non-homogeneous leukoplakia, erosive lichen planus, oral submucous fibrosis and actinic keratosis (Warnakulasuriya 2007).

The natural history of OSCC is not fully understood; not all PMD undergo malignant transformation and some affected sites can revert back to health (Napier 2008; Scully 2009). Equally, some OSCC can develop from lesions in which epithelial dysplasia was not previously diagnosed. Erythroplakia and erythro-leukoplakia have amongst the highest malignant transformation rates, followed by oral submucous fibrosis (Scully 2009). Oral leukoplakia is the most common PMD, but has a varied malignant

transformation rate (Reibul 2003; Napier 2008; Mehanna 2009; Scully 2009). Petti 2003 calculated a global malignant transformation rate (MTR) of oral leukoplakia of 1.36% per year (95% confidence interval 0.69 to 2.03%), but when this is applied to the prevalence of the condition, it far exceeds the numbers of actual cases of malignancy that present. However, the MTR in hospital-based studies is consistently higher than the MTR in community-based studies.

The early detection and excision of PMD can prevent malignant transformation (Warnakulasuriya 2007; van der Waal 2009; Brocklehurst 2013). 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 2008). Systematic reviews have identified no experimental evidence for surgical interventions (including laser therapy and cryotherapy) and little experimental evidence for non-surgical interventions. In addition, where clinical resolution was observed, relapses were common (Lodi 2008; Mehanna 2009).

In the United Kingdom, patients presenting with oral lesions per-

sisting for more than two to three weeks are generally referred to Oral Medicine Units or Oral and Maxillofacial Surgery Units for further investigation (Scully 2000a; Scully 2000b; Scully 2000c). Technologies to treat and manage oral cancer have progressed substantially (Glenny 2010; Furness 2011; Bessell 2011). Despite this, mortality rates have remained high (approximately 50%) and have not improved over the last 30 years (Parkin 2001; Warnakulasuriya 2009). If the lesion has progressed to frank malignancy, 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 dental attendance and dental professional 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). Yet oral cancer can be cured if caught early enough (Stell 1982). It has been estimated that in the UK, 80% of the population will visit a dentist at least once in the previous five years (Tilley 2005). Hence the dental team must screen every patient they see, particularly irregular attenders.

Index test(s)

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

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

We evaluated all four categories of index tests in this review (restricting vital staining index tests 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 (Table 1), to aid in the diagnosis of OSCC and PMD.

Clinical pathway

There is no internationally recognised or standardised clinical pathway for individuals presenting with PMD. 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 PMD or malignancy is suspected, the frontline clinician refers on to an oral specialist for a definitive diagnosis and scalpel biopsy, as appropriate.

Supplementing the COE with an index test could aid in the identification of clinically evident lesions. Tests could have a triage role in assisting the general dentist or oral specialist to more accurately assess oral lesions of uncertain significance. It could also help to reduce the unnecessary referral of benign conditions. For instance, traumatic keratoses are common benign white lesions and referring every patient would be excessive and incur increased financial cost and anxiety for the referred patient. 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 have the potential to improve patient diagnosis at a secondary care level. Following referral to a specialist clinic, a scalpel biopsy is commonly undertaken on areas representing the worst of the disease. This becomes complicated when the lesion or lesions under investigation are large or heterogeneous in nature. Sample site selection may be facilitated by the use of diagnostic adjuncts. The tests could also be used to help monitor a patient who has had a history of oral cancer or PMD. This population has often been exposed to surgical procedures, radiotherapy and repeated biopsies. Monitoring these patients for new disease can become complicated due to field change, where previously healthy mucosa undergoes malignant transformation. Determining the most appropriate site for a biopsy is challenging and the diagnostic adjuncts could be of value.

Alternative test(s)

The COE has been investigated as a test for early detection of PMD (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 incisional biopsy (which acts as the reference standard).

Rationale

Oral cancer is a significant global health problem with increasing incidence and mortality rates (Warnakulasuriya 2009; Ferlay 2010). Cancer of the lip or oral cavity is a relatively common cancer worldwide, with an estimated 263,000 new cases and 127,000 deaths in 2008, and an increasing incidence of disease in recent years (Ferlay 2010). There is wide geographic variation in disease incidence and mortality, with almost double the incidence in developing countries compared to developed countries, and a three-fold increase in mortality. Tobacco use, alcohol consumption, betel quid chewing and low socioeconomic status are the most important risk factors for oral cancer (Macfarlane 1995; Faggiano

1997; La Vecchia 1997; Conway 2008). 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, the incidence in recent years among 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 (Scully 2000a). Accurate case detection and early treatment of oral cancers can substantially improve an individual's morbidity, mortality and quality of life (Stell 1982; Scully 2000a). However, no national population-based screening programmes for oral cancer have been implemented in developed countries, although opportunistic screening has been advocated (Brocklehurst 2013). A province-wide programme is being evaluated in British Columbia, Canada but the evaluation is ongoing and no final results have been reported to date (Rosin 2006). Brocklehurst et al's Cochrane systematic review identified one randomised controlled trial (RCT) from India. They concluded that the evidence is insufficient to recommend population-based screening but opportunistic screening of high-risk groups may potentially improve outcomes (Brocklehurst 2013).

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; Furness 2011), but the five-year survival following diagnosis has remained constant at around 50% for the past 30 years (Parkin 2001; Warnakulasuriya 2009; Cancer Research UK).

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

This diagnostic test accuracy review complements a number of intervention reviews undertaken by the Cochrane Oral Health Group on the treatment of oral and oropharyngeal cancers (Glenny 2010; Furness 2011; Bessell 2011), 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 potentially malignant disorders in apparently healthy adults (Walsh 2013).

OBJECTIVES

To estimate the diagnostic accuracy of index tests for the detection of oral cancer and potentially malignant disorders of the lip and oral cavity, in patients presenting with clinically evident 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 lesions. The index tests assessed in this review are provided in 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 reported in abstract form alone, case-control studies, uncontrolled reports and randomised controlled trials (RCTs) of the effectiveness of screening programmes (intervention studies).

We had intended to exclude studies that analysed at the lesion level rather than at the individual level, unless the authors could provide patient-level data. In a change from the original protocol, we included studies reporting at the lesion level, and identified these studies in any analyses.

Participants

Adults (aged 16 years or over) presenting with clinically evident oral lesions.

Index tests

Index tests used alone, or in combination, as an adjunct to the COE (Table 1). Where multiple index tests were used together, we classed as positive a positive test result from the COE or the index test or both.

Target conditions

Following the consensus views of the expert working group of the WHO collaborating centre for oral cancer and pre-cancer (Warnakulasuriya 2007), the following target conditions of the lip or oral cavity were considered for inclusion in this review.

Carcinoma

- Oral squamous cell carcinoma (OSCC)

Potentially malignant disorders (PMD)

- Leukoplakia
- Erythroplakia
- Lichen planus

- Lupus erythematosus
- Submucous fibrosis
- Actinic keratosis
- Hereditary disorders such as dyskeratosis congenita or epidermolysis bullosa

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

Reference standards

Scalpel, punch or fine needle aspiration biopsy with histological diagnosis. Studies not specifying any reference standard were ineligible for inclusion in this review.

Search methods for identification of studies

Electronic searches

We searched the following electronic databases.

- Cochrane Oral Health Group's Trials Register (to 30 April 2013) (see [Appendix 1](#))
- Cochrane Register of Diagnostic Test Accuracy Studies (to 30 April 2013) (see [Appendix 2](#))
- MEDLINE via OVID (1946 to 30 April 2013) (see [Appendix 3](#))
- EMBASE via OVID (1980 to 30 April 2013) (see [Appendix 4](#))
- MEDION (2003 to 30 April 2013) (see [Appendix 5](#))

The MEDLINE search strategy outlined in [Appendix 3](#), was modified for the listed databases and was based on the companion Cochrane diagnostic test accuracy review ([Walsh 2013](#)) undertaken by the same team. The search was not limited by language or publication status. Non-English studies were translated.

Searching other resources

We searched the reference lists of key studies and contacted authors of studies to ask if they were aware of any unpublished or ongoing studies.

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, socioeconomic 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 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 ([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 a 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

We entered data for the true positive, true negative, false positive and false negative values for each test in each study into Review Manager (Rev Man 2014). The unit of analysis was the lesion. Where study results were only recorded at the lesion level, these studies were included as reported but identified in the forest plots. The average number of lesions per individual analysed is provided in the Characteristics of included studies section. For each index test, estimates of the diagnostic accuracy are expressed as sensitivity and specificity along with 95% confidence intervals. This information is displayed as coupled forest plots, and plotted in receiver operating characteristic (ROC) space. 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 (Reitsma 2005).

Indirect pairwise analyses were also structured as follows:

- vital staining versus brush cytology;
- light detection versus brush cytology;
- vital staining versus light detection.

We included all studies in each pairwise comparison. We also presented results of studies that directly compared different index tests (i.e. paired data from all individuals or individuals randomised to different index tests) in additional forest plots. SAS and Stata (version 13) were used for all statistical analyses. In light of the number of studies available, we also undertook comparative analyses where we relaxed the assumption of equal variances.

Investigations of heterogeneity

We carried out meta-regression analyses to explore possible sources of heterogeneity. Covariates in these analyses included:

- characteristics of the study sample: prevalence of the disease in the study;
- methods used by index test: rinsing or targeted staining with the vital staining products or classification of fluorescence device for light-based studies.

We had intended to investigate the severity of the target condition as another potential source of heterogeneity. As the majority of

studies included 'any dysplasia' as the target condition, there was insufficient variation in the included studies to do this.

The log likelihood of models including the covariate was compared to those models without the covariate. Formal model comparisons were undertaken using the likelihood ratio statistic to statistically compare the effects of adding or removing one or more covariate, with the assumption of equal variances.

Sensitivity analyses

We intended to carry out a sensitivity analyses that restricted the analysis to studies where the reference standard was scalpel biopsy followed by histopathology. We were unable to do this as the reference standard in all the included studies was scalpel biopsy.

Assessment of reporting bias

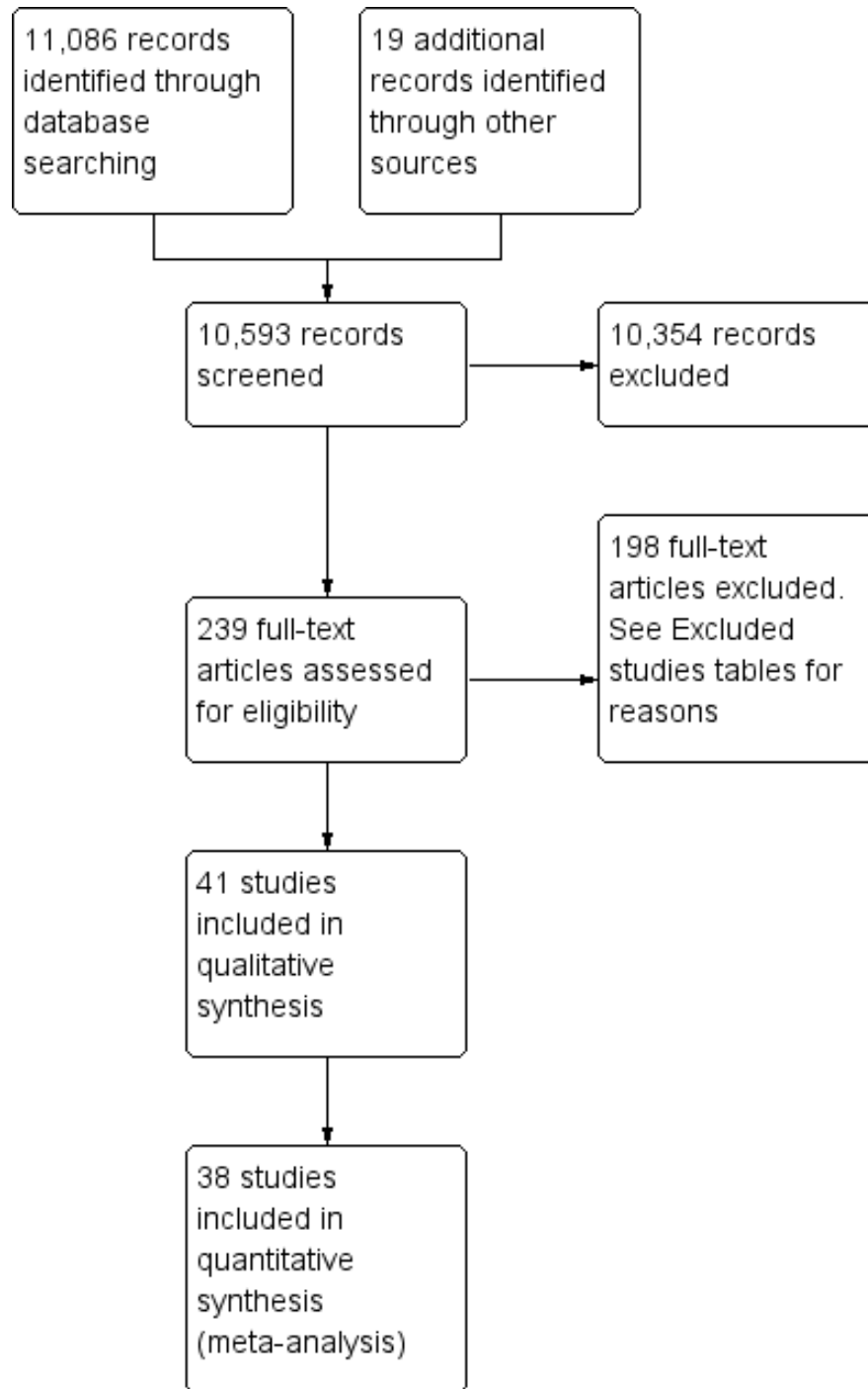
We did not undertake tests for reporting bias as these can be misleading when applied to systematic reviews of diagnostic test accuracy (Tang 2000; Leeftang 2008).

RESULTS

Results of the search

The search identified a total of 10,593 records, after duplicates were removed. We excluded 10,354 in accordance with the eligibility criteria and the remaining 239 studies were assessed for eligibility. Forty-one of these studies were eligible for inclusion (Figure 1). These studies evaluated data from a total of 4002 patients and 4337 lesions, with a mean number of patients per study of 97.61. There was a broad geographical spread: 18 of the studies were of European origin, 12 from India, five from the US, three from Australia and one from Brazil; the remainder originated in Asia. The studies were conducted between 1980 and 2012.

Figure 1. Study flow diagram



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 OSCC or PMD. All included studies were carried out in a secondary care (hospital) setting. Thirty studies evaluated a single adjunct test on a single sample. Of the remaining 11 studies, four assessed multiple tests on the same sample (Mehrotra 2010; Awan 2011; Ujaoney 2012; Rahman 2012); one assessed a single test on different samples (Cheng 2003a; Cheng 2003b) and six evaluated index tests used in combination (Gupta 2007; Epstein 2008; Mehrotra 2010; Guneri 2011; Mojsa 2012; Ujaoney 2012).

For the purposes of evaluation:

- Fourteen studies evaluated vital staining: 12 used toluidine blue (or methylene blue), one used toluidine blue and/or Lugol's Iodine (Nagaraju 2010) and one used Rose Bengal (Du 2007). Nine of these studies used a mouth rinse technique whilst four used a direct staining approach. One study (Cheng 2003a; Cheng 2003b) made a comparison between rinse and stain and has therefore been reported separately to allow for analysis of this covariate.
- Thirteen studies evaluated oral cytology, 12 of which used a brush to harvest cells: five assessed the Oral CDx system (Sciubba 1999; Svirsky 2002; Scheifele 2004; Delavarian 2010; Seijas-Naya 2012), three assessed the Cytobrush (Navone 2004; Koch 2011a; Rahman 2012) and the remainder used other

brush-based methods (Mehrotra 2008; Remmerbach 2009; Mehrotra 2011; Ng 2012). One study (Navone 2008) used a curette to harvest cells.

- Thirteen studies evaluated light-based technologies: four assessed the VELScope device (Mehrotra 2010; Awan 2011; Scheer 2011; Farah 2012), three assessed the Vizilite device (Farah 2007; Awan 2011; Ujaoney 2012); the remaining six assessed other light-based technologies.

- Six studies evaluated index tests used in combination: four assessed vital staining with light-based detection (Epstein 2008; Mehrotra 2010; Mojsa 2012; Ujaoney 2012) and two assessed the use of vital staining with cytology (Gupta 2007; Guneri 2011).

All studies used a reference test of scalpel biopsy and histological examination; no punch or fine needle biopsies were observed. All studies were set within a secondary care environment.

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. No single study could be classified as being at low risk of bias across all domains. 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

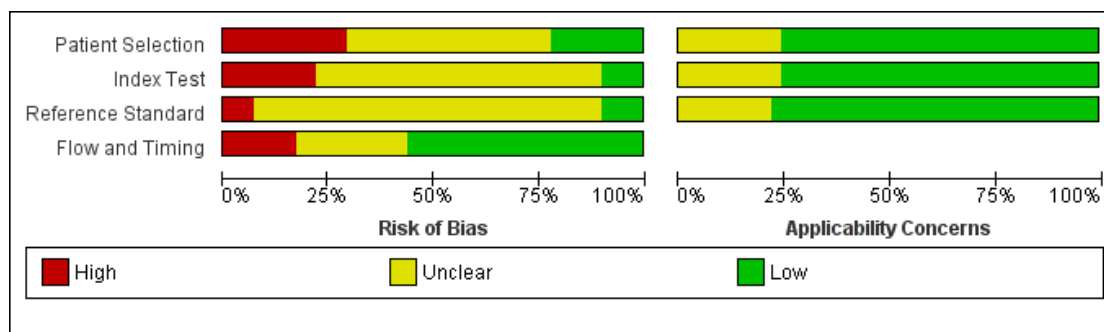
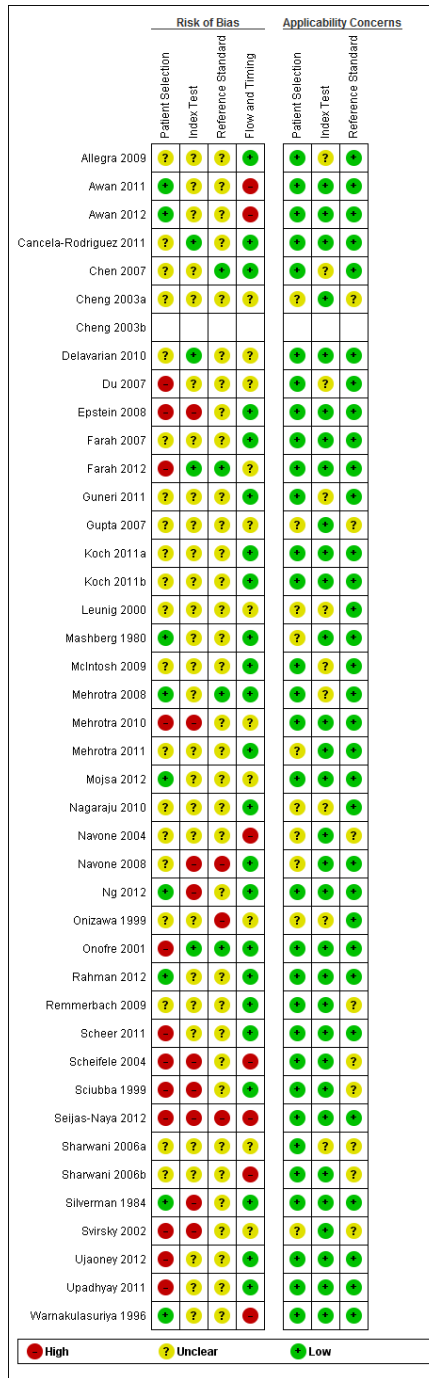


Figure 3. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study



Patient selection was considered to be at low risk of bias in nine out of 41 of studies (22%). Twenty of the 41 studies (49%) failed to describe the patient selection criteria in sufficient detail and were therefore assessed as being at unclear risk of bias. Eleven studies selected patients that were either at high or low risk and therefore inappropriately excluded people (Sciubba 1999; Svirsky 2002; Scheifele 2004; Du 2007; Epstein 2008; Mehrotra 2010; Scheer 2011; Upadhyay 2011; Seijas-Naya 2012; Farah 2012; Ujaoney 2012). Onofre 2001 failed to use consecutive or random sampling. The index test was considered to be at unclear risk of bias in 28 out of 41 studies (68%). The most common reason was a lack of detail about training or calibration of the clinicians. Only six of the included studies had a training or calibration component (Sciubba 1999; Onofre 2001; Mehrotra 2011; Awan 2011; Cancela-Rodriguez 2011; Farah 2012). In addition, a number of conflicts of interest were found where authors had links to companies supplying or developing the diagnostic aids being investigated (Silverman 1984; Sciubba 1999; Svirsky 2002; Scheifele 2004; Epstein 2008; Ng 2012). The positivity threshold was assessed as being at a high risk of bias in one study (Navone 2008). There was uncertainty in how the results of the cytology were graded in two studies, when there was an inadequate harvest of basal cells (Sciubba 1999; Scheifele 2004), which made interpretation and data extraction difficult. Finally, it was not possible to determine the order of the index tests in one study where multiple index tests were used (Mehrotra 2010).

All of the included studies used an appropriate reference standard; a scalpel biopsy followed by a histological examination by an experienced oral pathologist. However, many did not provide sufficient detail about the scalpel biopsy or the thresholds used in the histological examination. Three studies failed to ensure that the reference test was conducted independent of the results of the index test: Onizawa 1999, Navone 2008 and Seijas-Naya 2012, in which the same pathologist undertook the brush biopsy and the histological examination.

We deemed 23 of the studies to be at low risk of bias regarding the flow and timing of the study. 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. We classified 11 studies as unclear as they failed to report the tim-

ing difference. We found seven studies to be at high risk, five of these excluded participants from the analysis despite them receiving both index and reference tests (Warnakulasuriya 1996; Navone 2004; Scheifele 2004; Sharwani 2006b; Awan 2012), while one reported a delay of three weeks between index and reference tests (Seijas-Naya 2012) and one allowed participants to receive the index test without a reference test (Awan 2011).

We assessed 20 studies as having low concern for applicability: 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 satisfy us of the applicability of the study methods.

Findings

Vital staining

From the 14 studies evaluating this test (Cheng 2003a evaluated two separate samples resulting in 15 data points for sensitivity and specificity), we included 1248 individuals and 1338 lesions in the analysis. The definition of the target condition varied between the studies: all classified OSCC as positive; 12 of the studies reported any dysplasia as positive or we were able to re-classify according to this threshold. Ujaoney 2012 used a positivity threshold of 'high risk lesion' and we were unable to re-analyse. Rahman 2012 was also unclear but an assumption was made that any blue staining was classified as positive. Three studies reported data at the lesion level (Mashberg 1980, average number of lesions per individual 1.32; Ujaoney 2012, 1.80; Warnakulasuriya 1996, 1.42).

The summary estimates obtained from the meta-analysis were: sensitivity of 0.84 (0.74 to 0.90) with specificity of 0.70 (0.59 to 0.79) (Figure 4; Figure 5). If vital staining was used to identify PMD or OSCC in a sample of 1000 people, of whom 500 had the target condition (a median prevalence of positive lesions from the histological diagnosis of vital staining studies of 50%), then the target condition would go undetected in 80 individuals (false negative result) and 150 individuals without disease would be misclassified with a positive result (false positive result).

Figure 4. Forest plot of Vital staining

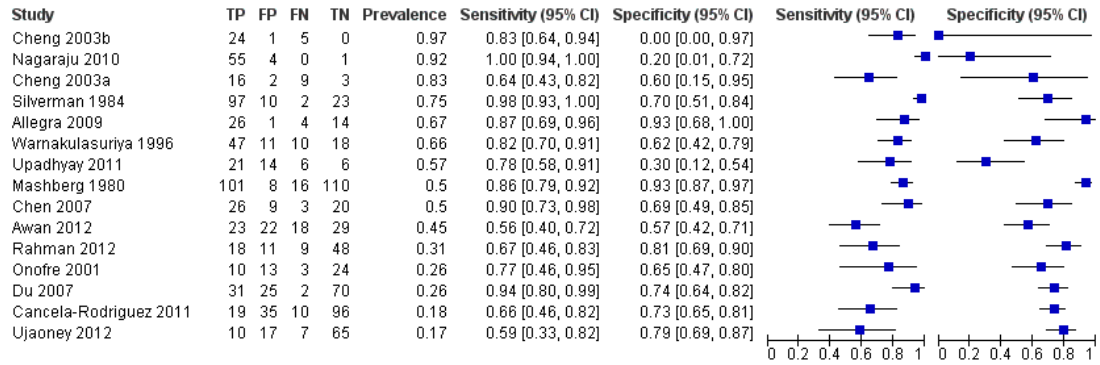
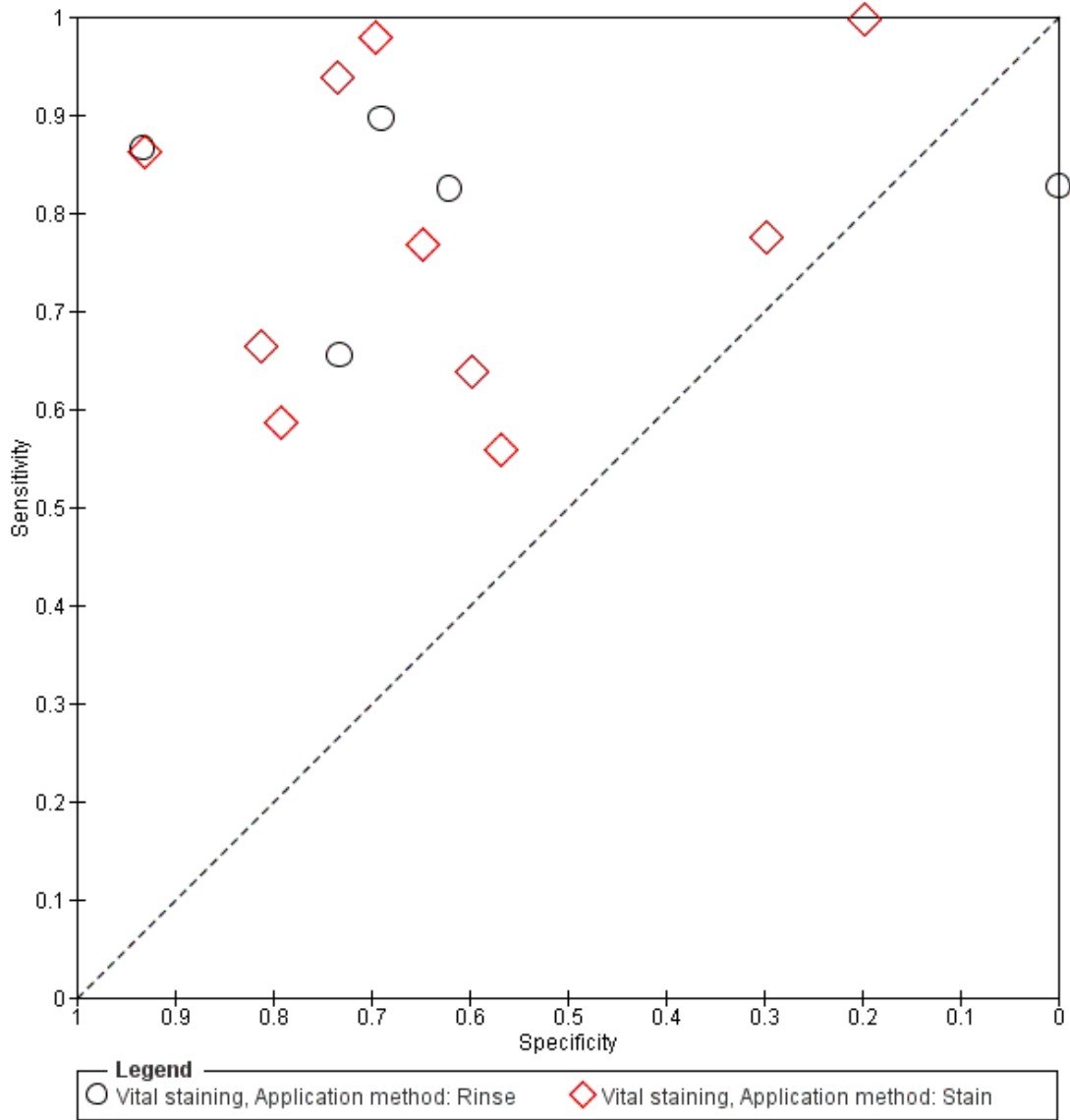


Figure 5. Summary ROC Plot of Vital staining



The coupled forest plot is presented along with the estimates of sensitivity and specificity for each study and plotted in ROC space. There was considerable variation in the estimates of both sensitivity and specificity. This is reflected in the coverage of the 95% confidence region (uncertainty in the overall average values of sensitivity and specificity due to sampling variation) and the 50% prediction region (a measure of between-study variability corresponding to the equivalent of an interquartile range in this instance).

We undertook meta-regression analysis to explore potential sources of heterogeneity as specified in the protocol. The results of the covariate analysis indicated that neither sensitivity nor specificity were associated with prevalence of disease ($P = 0.14$) or mode of administration (five studies used a rinse compared to 10 studies using a stain) ($P = 0.95$).

Oral cytology

In the 13 studies that evaluated oral cytology, 1554 participants and 1622 lesions were examined. The definition of the target con-

dition was consistently OSCC, carcinoma in situ (CIS) and all forms of dysplasia; except for [Rahman 2012](#), which created an “atypia” group and there was no clarity on how these were treated. Three studies reported data at the lesion level ([Delavarian 2010](#), average number of lesions per individual 1.04; [Koch 2011a](#), 1.35; [Scheifele 2004](#), 1.20).

We excluded one study that reported the results of multiple direct cytology tests ([Remmerbach 2009](#)) from the meta-analysis due to the potential of triple counting (as per guidance in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*, [Macaskill 2010](#)). The summary estimates obtained from the meta-analysis of 12 studies of sensitivity were 0.91 (0.81 to 0.96) and a specificity of 0.91 (0.81 to 0.95) ([Figure 6](#); [Figure 7](#)). If cytology was used to identify OSCC and PMD in a sample of 1000 people, with the cytology studies reporting a median prevalence of the target condition of 43%, then the disease would go undetected in 39 individuals and 51 individuals without disease would be misclassified with a positive result.

Figure 6. Forest plot of Cytology

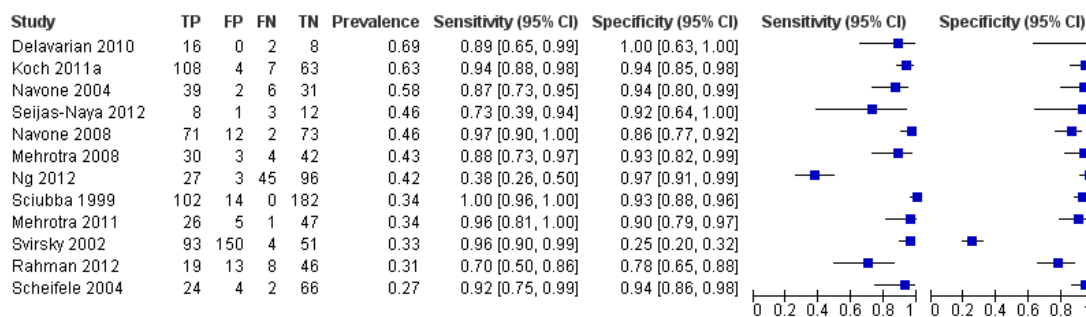
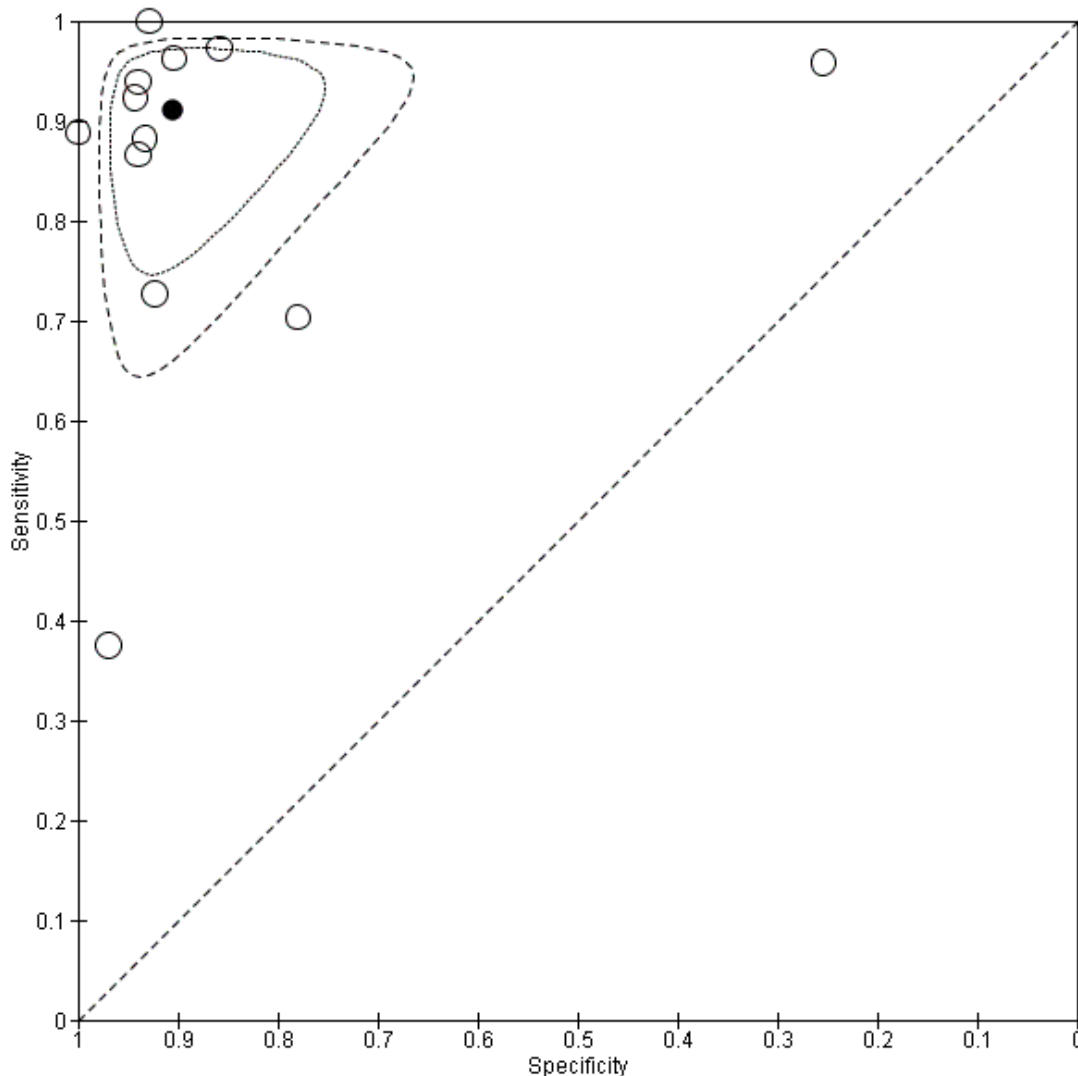


Figure 7. Summary ROC Plot of Cytology



The coupled forest plot is presented along with the estimates of sensitivity and specificity for each study and plotted in ROC space. There was some variation in the estimates of sensitivity; this was less evident for estimates of specificity. This is reflected in the coverage of the 95% confidence region and the 50% prediction region. There are two studies for which the results are atypical of the included studies. One study reported low values of specificity but high values of sensitivity (Svirsky 2002); one study reported low values of sensitivity but high values of specificity (Ng 2012). We undertook meta-regression analysis to explore the prevalence of dysplasia and OSCC as a potential source of heterogeneity

as specified in the protocol. The results of the covariate analysis showed that neither sensitivity nor specificity were associated with prevalence of disease ($P = 0.45$). We had planned to investigate different methods of biopsy, but there were too few studies for each method for the results of the analysis to be robust: OralCDX (five studies), CytoBrush (two studies), curette (one study) and other (six studies).

Light-based detection and/or oral spectroscopy

Thirteen studies evaluated data from 1253 patients and 1397 le-

sions. The target conditions were OSCC and all forms of dysplasia in 10 of the studies, and we were able to re-classify one other study according to this threshold (Leunig 2000). One study reported a definition of the target condition as “high-risk” (Ujaoney 2012) and another reported OSCC alone (Scheer 2011). The studies by Scheer 2011 and Ujaoney 2012 also excluded patients with advanced OSCC and therefore failed to report data from the full patient spectrum. Three studies reported data at the lesion level (Farah 2012, average number of lesions per individual 1.05; Leunig 2000, 2.76; Ujaoney 2012, 1.80).

We excluded one study from the meta-analysis due to incomplete data reporting (Kulapaditharon 1998). In this study the diagnostic test accuracy of white light endoscopy was evaluated in 11 patients, but true negative results were not fully reported and therefore no specificity estimates could be calculated. We also excluded one study that reported the results of multiple direct light-based tests (Awan 2011) from the meta-analysis due to the potential of double counting (as per guidance in Macaskill 2010). The 126 patients included in this study would make up a high proportion of

participants for this index test and therefore have an impact on the pooled estimate. The sensitivity and specificity of autofluorescence (VELScope) for the detection of a dysplastic lesion were 0.84 and 0.15, respectively; the sensitivity and specificity of chemiluminescence (Vizilite) for the detection of a dysplastic lesion were 0.77 and 0.28, respectively (Awan 2011).

The meta-analysis yielded estimates of sensitivity of 0.91 (0.77 to 0.97) and specificity of 0.58 (0.22 to 0.87) (Figure 8; Figure 9). If a light-based detection method was used to identify OSCC and PMD in a sample of 1000 people, with the light-based studies reporting a median prevalence of the target condition of 21%, then the disease would go undetected in 19 individuals and 332 individuals without disease would be mis-classified with a positive result. The coupled forest plot is presented along with the estimates of sensitivity and specificity and plotted in ROC space. There are two studies for which the results are atypical of the included studies. Both of these studies reported very low values of specificity with very high values of sensitivity (Farah 2007; Ujaoney 2012).

Figure 8. Forest plot of Light-based

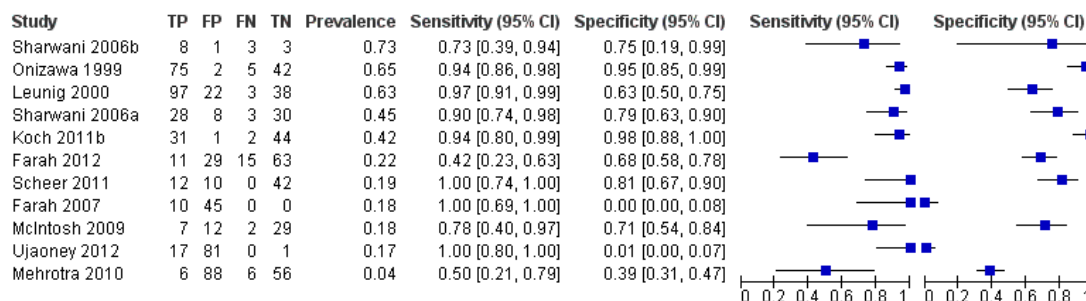
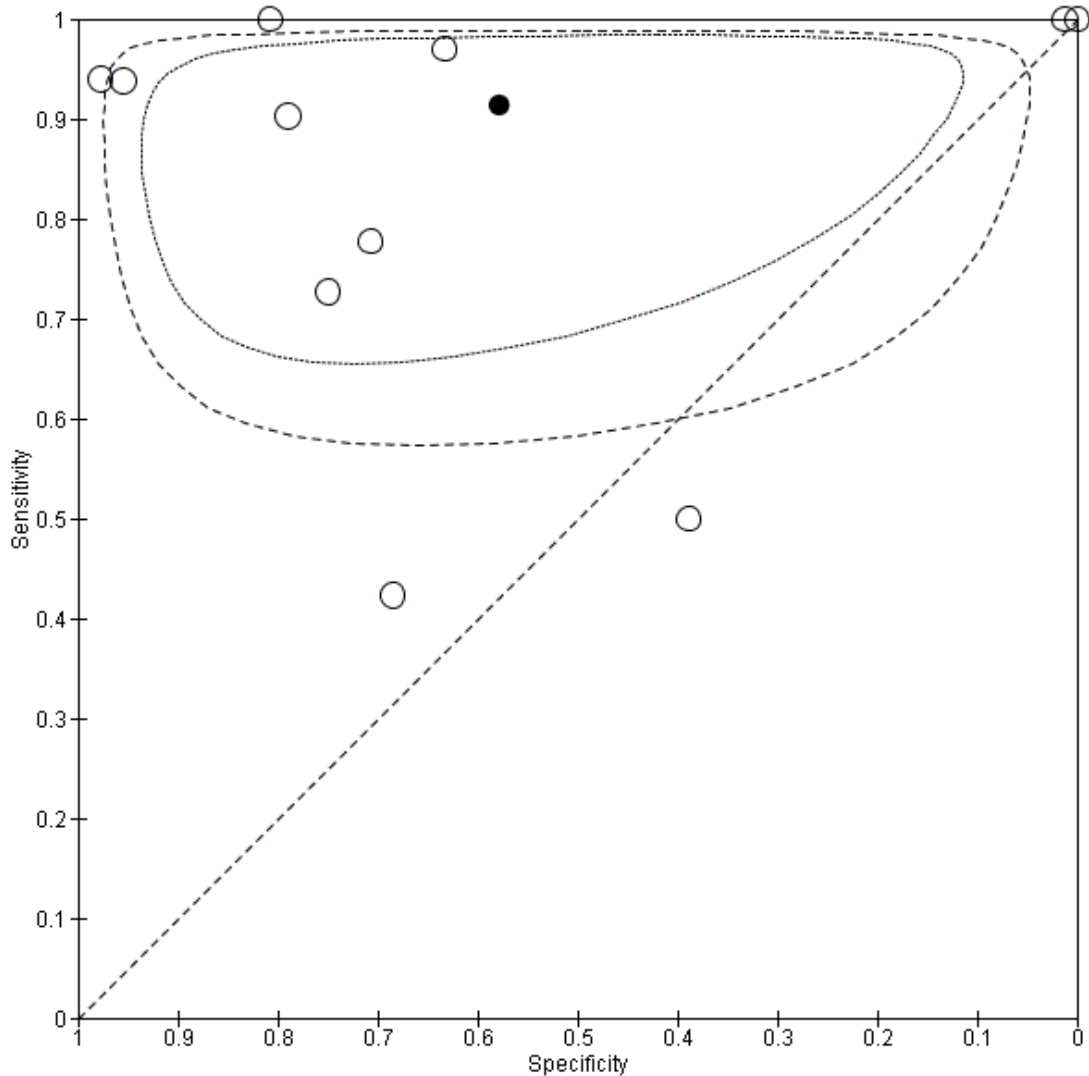


Figure 9. Summary ROC Plot of Light-based



Vital staining plus adjunct

Six studies with 402 individuals and 478 lesions used a vital staining with an adjunctive diagnostic test: four assessed vital staining with light-based detection (Epstein 2008; Mehrotra 2010; Mojsa 2012; Ujaoney 2012) (see Data table 4) and two assessed the use of vital staining with cytology (Gupta 2007; Guneri 2011) (see Data table 5). Due to the small number of studies, we did not carry out a meta-analysis of this data.

Relative performance of different tests

We added a covariate of test type to the bivariate analysis to ascertain the relative diagnostic test accuracy of the different tests. In accordance with the protocol, we carried out three pairwise comparisons. No difference in model fit was observed where equal variances were assumed except for the comparison highlighted below.

- Vital staining versus cytology
- Vital staining versus light-based detection (not assuming equal variances)

- Cytology versus light-based detection

The results of the model comparisons are presented in [Table 3](#).

Vital staining versus cytology

This comparison was based on data from 14 studies of vital staining (from 15 different samples) and 12 studies of cytology. The coupled forest plot is presented along with the estimates of sensitivity and specificity and plotted in ROC space. The analysis contained paired data for one study ([Rahman 2012](#)). Initial analysis indicated that the expected summary measures of sensitivity and/or specificity differed between the vital staining and cytology tests ($P = 0.004$). Further exploration indicated that there was no statistical evidence that sensitivity differed between vital staining compared to cytology ($P = 0.23$), but that there was strong statistical evidence that specificity differed between vital staining compared to cytology ($P = 0.003$). (Vital staining sensitivity 0.84 (0.74 to 0.90) and specificity 0.70 (0.59 to 0.79); cytology sensitivity 0.91 (0.81 to 0.96) and specificity of 0.91 (0.81 to 0.95)). There was no statistical evidence ($P = 0.55$) to suggest that the assumption of equal variances was violated. Assuming the variances to be the same the values for sensitivity and specificity of stain and cytology the following values resulted: Vital staining sensitivity 0.84 (0.73 to 0.91) and specificity 0.69 (0.55 to 0.80); cytology sensitivity 0.91 (0.82 to 0.96) and specificity of 0.90 (0.82 to 0.95).

Vital staining versus light-based detection

This comparison was based on data from 14 studies of vital staining (from 15 different samples) and 11 studies of light-based detection. The coupled forest plot is presented along with the estimates of sensitivity and specificity and plotted in ROC space.

The analysis contained paired data for one study ([Ujaoney 2012](#)). Results of the analyses indicated that the expected summary measures of sensitivity and/or specificity did not differ between the vital staining and light-based detection ($P = 0.49$). (Vital staining sensitivity 0.84 (0.74 to 0.90) and specificity 0.70 (0.59 to 0.79); light-based detection sensitivity of 0.91 (0.77 to 0.97), specificity 0.58 (0.22 to 0.87)). For this comparison there was statistical evidence ($P = 0.01$) to suggest that the assumption of equal variances may not be reasonable (estimates not assuming equal variances are provided).

Cytology versus light-based detection

This comparison was based on data from 12 studies of cytology and 11 studies of light-based detection. The coupled forest plot is presented along with the estimates of sensitivity and specificity and plotted in ROC space. Initial analysis indicated that there was weak evidence that expected summary measures of sensitivity and/or specificity differed between the cytology and light-based detection ($P = 0.06$). Further exploration indicated that there was no statistical evidence that sensitivity differed between cytology and light-based detection ($P = 0.99$), but that there was strong statistical evidence that specificity differed between vital staining compared to cytology ($P = 0.02$). (Cytology sensitivity 0.91 (0.81 to 0.96) and specificity of 0.91 (0.81 to 0.95), light-based detection sensitivity 0.91 (0.77 to 0.97) and specificity 0.58 (0.22 to 0.87)).

Comparison of the models with covariates assuming and not assuming equal variances was not statistically significant ($P = 0.21$). Assuming the variances to be the same the following values resulted: cytology sensitivity 0.91 (0.81 to 0.96) and specificity of 0.91 (0.77 to 0.97), light-based detection sensitivity 0.91 (0.78 to 0.97) and specificity 0.59 (0.31 to 0.82).

Summary of findings

What is the most accurate health technology for diagnosing oral cancer and potentially malignant disorders?					
Patient population	Patients referred to a secondary care facility for further investigation of a clinically evident lesion				
Index test	Conventional oral examination and adjunctive test (Vital Stain, Oral Cytology, Light-based)				
Reference test	Scalpel biopsy and histological assessment by experienced oral pathologist				
Target condition	Oral cavity cancer or potentially malignant disorder				
Included Studies	41 cross-sectional studies				
Test/Subgroup	Summary accuracy: Sensitivity (95% CI)	Summary accuracy: Specificity (95% CI)	No. of participants/ lesions/studies	Prevalence Median (range)	Quality and Comments
Vital Staining	0.84 (0.74 to 0.90)	0.70 (0.59 to 0.79)	1248 / 1338 / 14	0.50 (0.17 to 0.97)	High risk of bias ⁴
Oral Cytology	0.91 (0.81 to 0.96)	0.91 (0.81 to 0.95)	1507 / 1575 / 12 ¹	0.43 (0.27 to 0.69)	High risk of bias ⁵
Light-based	0.91 (0.77 to 0.97)	0.58 (0.22 to 0.87)	1021 / 1165 / 11 ²	0.21 (0.04 to 0.73)	High risk of bias ⁶
Vital Stain plus adjunct	Not possible ³	Not possible ³	402 / 478 / 6 ³	0.27 (0.04 to 0.48)	High risk of bias ⁷

¹ Included in the meta-analysis ([Remmerbach 2009](#) not included).

² Included in the meta-analysis ([Awan 2011](#) not included, which performed two index tests).

³ Insufficient studies to perform meta-analysis, data presented identifies numbers included in the review.

⁴ High risk of bias: inappropriate exclusion of patients ([Du 2007](#); [Upadhyay 2011](#)), selecting patients ([Onofre 2001](#)) and excluding patients from analysis ([Warnakulasuriya 1996](#)).

⁵ High risk of bias: inappropriate exclusion of patients ([Sciubba 1999](#); [Svirsky 2002](#); [Navone 2004](#); [Scheifele 2004](#); [Seijas-Naya 2012](#)), index and reference test conducted simultaneously ([Navone 2008](#)), conflict of interests ([Ng 2012](#); [Scheifele 2004](#); [Sciubba 1999](#); [Svirsky 2002](#)), time between index and reference test greater than two weeks ([Scheifele 2004](#); [Seijas-Naya 2012](#)).

⁶ High risk of bias: excluding patients from analysis ([Awan 2011](#); [Awan 2012](#); [Sharwani 2006b](#)), inappropriate exclusion of patients ([Farah 2012](#); [Scheer 2011](#)).

⁷ High risk of bias: inappropriate exclusion of patients ([Epstein 2008](#); [Mehrotra 2010](#); [Ujaoney 2012](#)), conflict of interests ([Epstein 2008](#)), multiple index tests not interpreted independently ([Mehrotra 2010](#)).

DISCUSSION

Summary of main results

Our review aimed to estimate the accuracy of index tests for the detection of oral cancer and potentially malignant disorders (PMDs) of the lip and oral cavity in patients presenting with clinically evident lesions. There were studies available to assess three diagnostic technologies, namely, vital staining, cytology and light-based. Despite all being undertaken in secondary care facilities, the studies showed diversity in geographic location, prevalence of disease, relative skill level of practitioners, outcomes assessed and thresholds used. However, a sufficient number of studies were identified to proceed with meta-analysis. The methodological quality of the studies was assessed to be poor overall, with no single study being at low risk of bias across all the domains of the risk of bias.

The estimates of sensitivity and specificity for cytology were the highest of the three health technologies assessed (summary sensitivity 0.91 and specificity 0.91). This overall high level of test accuracy requires careful interpretation. When expert clinicians encounter patients with PMDs they initiate a diagnostic algorithm, procure tissue and submit it for histopathologic evaluation to render a definitive diagnosis. This is the established gold standard for the diagnosis of oral epithelial dysplasia and oral squamous cell carcinoma (OSCC). Indeed, when such an algorithm involves clinicians with expertise in the diagnosis of oral mucosal diseases who are experienced in performing minimally invasive biopsies with optimal site selection, and is coupled with tissue interpretation by expert pathologists, it provides a very efficient pathway to an accurate diagnosis. Yet, unfortunately, such clinicians are generally not the frontline clinicians who initially encounter patients with PMDs. As such, the prism through which these adjunctive technologies are interpreted must be considered. The high accuracy of cytology we report is based on studies predominantly performed by such experts in cohorts with a higher percentage of dysplastic or malignant lesions than would be expected in a general population. Indeed, the strong correlation between the cytological samples that are procured from the same site(s) that are biopsied should not be a surprise. Yet, current cytologic test outcomes cannot discriminate between epithelial dysplasia or 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 for use 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, our review cannot adequately answer this question. The prevalence of PMDs in the general population is relatively high compared to the known incidence of OSCC and therefore most of

the PMDs 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. 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. Clinicians must appreciate that cytologic tests are not indicated for persistent epithelial lesions for which a clear aetiology is unapparent and which display one or more high-risk clinical features (e.g. induration, pain, ulceration and/or heterogeneous white, red or mixed red and white 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 surveillance 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.

Light-based technologies (summary sensitivity 0.91 and specificity 0.58) are hampered by their low specificity. Frontline clinicians relying on these technologies to dictate whether or not to recommend a biopsy or refer a patient for biopsy would result in more than 40% of patients receiving unnecessary referral/treatment. There is too much “noise” associated with these technologies that will confuse the non-expert clinician. Filtering the “noise” by gaining an appreciation for differentiating benign “confounder” lesions from PMDs (e.g. lichen planus, erythematous candidiasis, pigmented lesions and others) could bolster the specificity, but studies assessing these technologies in the hands of frontline clinicians are lacking.

Vital staining technologies (summary sensitivity 0.84 and specificity 0.70) are hampered by suboptimal sensitivity and specificity. The utility of vital staining should be confined for use by expert clinicians to facilitate biopsy site selection in heterogeneous PMDs, or as a surveillance tool.

The potential for vital staining and light-based detection to be used as a screening tool for detecting OSCC and PMD in apparently healthy individuals warrants further investigation (U.S. Preventive Services Task Force 2013; Walsh 2013), but, to date, none of these have actually been used as a screening intervention (Brocklehurst 2013) and the cost-effectiveness of using these methods over and above a visual screen would need to be justified. The concept of combining technologies to improve test accuracy seems reasonable; however, it is not possible to support the combining of such tests as the data from this review were limited; more studies are needed.

Ideally, the role of adjunctive tests is to reduce uncertainty in the diagnostic decision. With some tests this can be achieved by

exploring different threshold levels. However, this is not possible with any of these tests as they all dichotomise patients as either diseased or healthy. As a result, threshold analysis and area under the curve could not be investigated.

Strengths and weaknesses of the review

A strength of the review is the number of included studies (41) and the large amount of data that was evaluated (4314 lesions). This enabled a series of meta-analyses to be undertaken and, given the relative consistency in the classification of the target condition, the potential for an accurate estimation of summary points was considered to be high.

The review also identified a number of different categories of index test, each of which contained sufficient data to estimate summary values of sensitivity and specificity with an acceptable degree of precision. A further strength of this review was the ability to determine the relative diagnostic test accuracy between the different index tests. However, the prevalence across the subgroups varied from 4% to 97% and there was additional heterogeneity in the eligibility criteria used, which makes a direct comparison more problematic. For example, one study (Epstein 2008) included an oral examination as part of the recruitment of patients to the study, which may have introduced bias to the selection process by allowing the selection of high-risk patients who are potentially more straightforward to diagnose.

As the target condition can also 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 of the respective tests. In many of the vital staining and light-based studies, it was not clear how clinicians determined the most appropriate location of the biopsy. For 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 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 the address the review question. The study setting, patient selection, test conduct and interpretation were appropriate for the purposes of the review. The concerns raised were a result of inadequate reporting of the study index and reference test methods.

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

None of the adjunctive tests can be recommended as a replacement for the currently used standard of a scalpel biopsy and histological assessment. Yet, the performance of cytology compared to histopathology shows promise; however, there is insufficient 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 PMD, 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). 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 correctly harvesting basal cells from the oral mucosa.

Implications for research

The overall quality of the included studies was poor, so there is a need for further standardisation of research in this area to reduce bias. Diagnostic test accuracy (DTA) studies addressing this area of research would be welcomed, particularly those that report on a patient basis and that follow the STARD checklist (Bossuyt 2003) for reporting of diagnostic test accuracy studies. All new DTA studies should ensure that they fully address the domains within the QUADAS-2 tool (Whiting 2011). Patients should be recruited consecutively or randomly prior to any oral examination to avoid the potential for selection bias; training and calibration of examiners should be clearly explained; the methods and timings of the reference standard should be clearly reported; and an accepted classification of the target condition used. The numbers of true positives, true negatives, false positives and false negatives should also be clearly reported. The positivity threshold that has been most consistently applied throughout these studies 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.

Given the relatively high values of the summary estimates of sensitivity and specificity for cytology, this would appear to offer the most potential. Combined adjunctive tests involving cytology would warrant further investigation, only two studies investigated vital staining plus cytology (Gupta 2007; Guneri 2011) 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: benign, mild, moderate, severe dysplasia and OSCC. This would allow

further analysis to be undertaken to assess the severity of any misdiagnosis; if an OCSS is classified as a mild dysplasia by the index test then the consequences are significantly more severe than if a moderate dysplasia be classified as mild.

This review revealed some novel cytologic platforms. Remmerbach 2009 demonstrated the potential for new analysis methods such as AgNOR, while Ng 2012 reported the use of quantitative DNA cytology to maximise cytology. As the complexities of carcinogenesis are revealed through a better understanding of genetic and epigenetic alterations in PMDs, 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 diagnostic potential of blood and saliva analysis should also be investigated. Despite preliminary work being completed in this

area (Li 2004; Park 2009), study designs have so far been restricted to case-control studies.

ACKNOWLEDGEMENTS

We would like to thank Anne Littlewood for her advice on searching the literature; Derek Richards for his thoughtful feedback on the protocol; Helen Worthington and Jo Weldon for their assistance with data extraction; Laura MacDonald, Ruth Floate, Miguel Ángel González Moles, Rachel Hall and the Cochrane Diagnostic Test Accuracy Group who provided feedback and assisted with the editorial process; Giovanni Lodi, Anette Blümle, Joanna Zakrzewska, Chunjie Li and Janet Lear for the translation of papers; NHS Education Scotland and the Scottish Dental Clinical Effectiveness Programme for their support on this review; .

REFERENCES

References to studies included in this review

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 precancerose e cancerose dell'orofaringe e del cavo orale]. *ACTA Otorhinolaryngologica Italica* 2009;**29**:187–90.

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 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.

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 pre-malignant and malignant oral lesions. *Journal of Oral Pathology & Medicine* 2010;**40**:300–4.

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.

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;**Sep 1**(15):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.
- 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.
- 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.
- Mashberg 1980** *{published data only}*
Mashberg A. Reevaluation of toluidine blue application as a diagnostic adjunct in the detection of asymptomatic oral squamous carcinoma: a continuing prospective study of oral cancer III. *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, Shukla P. 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.
- 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, Maticic 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**:206–210.
- 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* 2001; **May**: 535–40.
- 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.
- 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; **June**:228–35.
- 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 Radiation and Endodontology* 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 Patología Oral y Cirugía 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-aminolevulinic acid-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.
- Silverman 1984** *{published data only}*
Silverman S, Migliorati C, Barbosa J. Toluidine blue staining in the detection of oral precancerous and malignant lesions. *Oral Medicine* 1984;**April**:379–82.
- 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.
- Ujaoney 2012** *{published data only}*
Ujaoney S, Morwani 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.
- References to studies excluded from this review**
- 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 OralBrush Biopsy” by Hohlweg-Majert et al. *Cancer Investigation* 2010;**28**(5):560–1.
- 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.
- Ebenazar 2012** *{published data only}*
Ebenazar J, Ganesan S, Aruna P, Muralinaidu R, Renganathan K, Saraswathy TR. Noninvasive 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, Maticic 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 toldina en enjuagues, como auxiliar diagnóstico de cancer orofaríngeo]. *Revista ADM* 1989;**46**(1):29–35.
- Hedge 2006** *{published data only}*
Hedge MC, Kamath PM, Shreedharan S, Dannana NK, Raju RM. Supravital staining: its role in detecting early malignancies. *Indian Journal of Otolaryngology and Head and 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, Boonkitticharoen V. Laser-induced fluorescence imaging in localization of head and neck cancers. *Annals of Otolaryngology & 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, Eckardt 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.

Reboiras-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. Noninvasive brush biopsy as an innovative tool for early detection of oral carcinomas [Nichtinvasive Bürstenbiopsie als innovative Methode in 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 [Minimalinvasive Bürstenbiopsie]. *Praxis Fortbildung* 2007;**117**:926–40.

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 RA1, Gao W, Redden Weber C, Kurachi C, Lee JJ, El-Naggar AK, et al. Noninvasive evaluation of oral lesions using depth-sensitive optical spectroscopy. *Cancer* 2009 April;**115**(8):1669–79.

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.

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.

Wang 2009 *{published data only}*

Wang CY, Tsai T, Chiang CB, 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.

Additional references**Bessell 2011**

Bessell A, Glenny AM, Furness S, Clarkson JE, Oliver R, Conway DI, et al. Interventions for the treatment of oral and oropharyngeal cancers: surgical treatment. *Cochrane Database of Systematic Reviews* 2011, Issue 9. [DOI: 10.1002/14651858.CD006205.pub3]

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.

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. [DOI: 10.1002/14651858.CD004150.pub3]

Buchen 2011

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

Cancer Research UK

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

Conway 2008

Conway DI, Petticrew M, Marlborough H, Berthiller J, Hashibe M, Macpherson LM. Socioeconomic inequalities and oral cancer risk: a systematic review and meta-analysis of case-control studies. *International Journal of Cancer* 2008;**122**(12):2811–9.

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.

Faggiano 1997

Faggiano F, Partanen T, Kogevinas M, Boffetta P. Socioeconomic differences in cancer incidence and mortality. In: Kogevinas M, Pearce N, Susser M, Boffetta P editor(s). *Social Inequalities and Cancer*. Lyon: IARC Scientific Publications No 138. International Agency for Research in Cancer, 1997.

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.

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. [DOI: 10.1002/14651858.CD006386.pub3]

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.

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. [DOI: 10.1002/14651858.CD006387.pub2]

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.

La Vecchia 1997

La Vecchia C, Tavani A, Franceschi S, Levi F, Corrao G, Negri E. Epidemiology and prevention of oral cancer. *Oral Oncology* 1997;**33**(5):302–12.

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.
- 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.
- 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.
- Lodi 2008**
Lodi G, Sardella A, Bez C, Demarosi F, Carrassi A. Interventions for treating oral leukoplakia. *Cochrane Database of Systematic Reviews* 2006, Issue 4. [DOI: 10.1002/14651858.CD001829.pub3]
- Macaskill 2010**
Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C editor(s). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0*. The Cochrane Collaboration, 2010. Available from www.srdta.cochrane.org/handbook-dta-reviews.
- Macfarlane 1995**
Macfarlane GJ, Zheng T, Marshall JR, Boffetta P, Niu S, Brasure J, et al. Alcohol, tobacco, diet and the risk of oral cancer: a pooled analysis of three case-control studies. *European Journal of Cancer. Part B, Oral Oncology* 1995;**31B**(3):181–7.
- 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.
- Napier 2008**
Napier SS, Speight PM. Natural history of potentially malignant oral lesions and conditions: an overview of the literature. *Journal of Oral Pathology and Medicine* 2008;**37**(1):1–10.
- 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. *The 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.
- Reibul 2003**
Reibul J. Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics. *Critical Reviews in Oral Biology and Medicine* 2003;**14**(1):47–62.
- 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.
- Rev Man 2014 [Computer program]**
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
- Rosin 2006**
Rosin M. Bringing the Promise of Molecular Medicine to Oral Cancer Screening. www.nidcr.nih.gov/Research/ResearchResults/InterviewsOHR/TIS032006.htm (accessed 18 April 2011).
- 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.
- Scully 2000a**
Scully C, Porter S. ABC of oral health. Oral cancer. *BMJ* 2000;**321**(7253):97–100.
- Scully 2000b**
Scully C, Shotts R. ABC of oral health. Mouth ulcers and other causes of orofacial soreness and pain. *BMJ* 2000;**321**(7254):162–5.
- Scully 2000c**
Scully C, Porter S. ABC of oral health. Swellings and red, white, and pigmented lesions. *BMJ* 2000;**321**(7255):225–8.

Scully 2009

Scully C, Bagan J. Oral squamous cell carcinoma overview. *Oral Oncology* 2009;**45**(4-5):301–8.

Stell 1982

Stell PM, Wood GD, Scott MH. Early oral cancer: treatment by biopsy excision. *British Journal of Oral Surgery* 1982;**20**(4):234–8.

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.

Tilley 2005

Tilley CJ, Chalkley MJ. Measuring access to health services: General Dental Services in Scotland. *British Dental Journal* 2005;**199**:599–601.

U.S. Preventive Services Task Force 2013

U.S. Preventive Services Task Force. Screening for Oral Cancer. <http://www.uspreventiveservicestaskforce.org/uspstf/uspsoal.htm> 2013.

van der Waal 2009

van der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa: terminology, classification and present concepts of management. *Oral Oncology* 2009;**45**(4-5):317–23.

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. [DOI: 10.1002/14651858.CD010173]

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.

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.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

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

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 (range 42-82) 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." 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: TB 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

Allegra 2009 (Continued)

	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		
Comparative			
Notes			
Methodological quality			
Item	Authors’ judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	No		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			

				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			
				Low
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			

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 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%) Alcohol: Current user 92 (73%); Ex-user 8 (6.4%); Never used 26 (20.6%) Number of patients/lesions: 126/126, 10 did not received reference test

	<p>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 two hospitals Prevalance: Dysplastic: 44/116 (VELScope & Vizilite),</p>
<p>Index tests</p>	<p>Index test 1 - VELScope Category: Light based Description: The main lesion was selected and diagnosed by consensus of two 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 blind 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</p> <p>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 blind 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.” Conflict of interests: Equipment provided by manufacturers and marketers of ViziLite quote “ViziLite kits for the study were supplied by Zila Inc.”</p>
<p>Target condition and reference standard(s)</p>	<p>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</p>

	<p>recorded by an experienced oral pathologist”</p> <p>Sequence of tests: Index tests followed by reference standard</p> <p>Training or calibration of pathologists: Only one experienced pathologist.</p> <p>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</p> <p>Multiple tests: No</p> <p>Method of site selection: The index test results were used to inform the selection of biopsy site</p> <p>Target condition: Dysplasia</p>
Flow and timing	<p>Patients receiving index test but not reference test: 10</p> <p>Patients receiving reference test but not index test: 0</p> <p>Time interval: the tests were conducted independently (though on the same session for patient convenience)</p> <p>Patients receiving both index and reference test but excluded from analysis: 0</p>
Comparative	
Notes	<p>Index test - VELScope & Vizilite: Only able to report on 116 patients that received reference test</p> <p>Papers report data on leukoplakia and erythroplakia (through clinical diagnosis only), only data reporting dysplastic or non-dysplasia from pathology used</p> <p>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</p>

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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		

Awan 2011 (Continued)

Was calibration of examiners undertaken and results reported?	No			
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?	Unclear			
				Low
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			
				Low
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			

Awan 2012

Study characteristics	
Patient sampling	Method of patient selection: Consecutive patients presenting in oral medicine clinics at two London Hospitals, with white, red and mixed white and red patches
Patient characteristics and setting	Setting: Secondary care, oral medicine clinics Age: Over 16 Sex: Male 61%; Female 39%

	<p>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 two hospitals Prevalance: 41/92</p>
Index tests	<p>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 tube 1, 1% acetic acid solution (pre rinse swab) , Swab tube 2, 0.5% toluidine 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 two examiners Conflict of interests: Equipment provided by manufacturers and marketers of Toluidine Blue kits, Zila Inc</p>
Target condition and reference standard(s)	<p>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 two experienced oral pathologists” Sequence of tests: Index tests followed by reference standard Training or calibration of pathologists: Two 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</p>
Flow and timing	<p>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</p>
Comparative	

Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?	Unclear		
			Low
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		
			Low

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

Cancela-Rodriguez 2011

Study characteristics	
Patient sampling	Method of patient selection: "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 13-100 years Sex: 77 men 83 women 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 pre-malignant or malignant lesions that were either white or red, exophytic or presenting as non-healing ulcers" Country: Spain Type of facility: Secondary Prevalance: 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: Quote "a working clinical diagnosis was established using WHO Criteria 1980", quote "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.		
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			
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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		

Cancela-Rodriguez 2011 (Continued)

Was calibration of examiners undertaken and results reported?	Yes			
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?				
				Low
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			
				Low
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			

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

	<p>Ethnicity: not reported</p> <p>Stated Risk Factors: quote "About two-thirds of cases consisted of smokers or people who chewed betel nuts"</p> <p>Number of patients/lesions: 58 patients and lesions</p> <p>Lesion site: Oral cavity</p> <p>Severity: quote "The lesions were characterised as homogeneous leukoplakia, heterogeneous leukoplakia, erythroplakia, and ulceration"</p> <p>Country: Taiwan</p> <p>Type of facility: Secondary</p> <p>Prevalance: (cancer or precancerous) 29/58</p>
Index tests	<p>Category: Vital staining - methylene blue</p> <p>Description: Quote "A standard staining procedure included gargling with rinsing solution (1% lactic acid in purified water flavoured with raspberry) for 20s, 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 20s followed by an additional 20s of gargling with the rinsing solution."</p> <p>Positivity threshold: unclear</p> <p>Sequence of tests: Index followed by reference</p> <p>Training or calibration of clinicians: Not reported. Quote "The pattern and intensity of the staining was recorded and supervised by an experienced oral surgeon."</p> <p>Blinding of examiners: Index completed prior to reference.</p> <p>Multiple tests: No</p> <p>Method of site selection: Methylene rinse so no site selection</p> <p>Conflict of interests: Quote "grant supported by NSC-94-2314B075 and VGH94242C"</p>
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: Biopsy method unclear</p> <p>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."</p> <p>Sequence of tests: Index followed by reference</p> <p>Training or calibration of pathologists: Quote "The pathological diagnosis was examined and verified independently by two specialists" Unclear</p> <p>Blinding of examiners: Examinations were independent</p> <p>Multiple tests: No</p> <p>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."</p> <p>Target condition: Dysplasia and squamous cell carcinomas</p>
Flow and timing	<p>Patients receiving index test but not reference test: 0</p> <p>Patients receiving reference test but not index test: 0</p> <p>Time interval: Completed at the same time</p> <p>Patients receiving both index and reference test but excluded from analysis: 0</p>
Comparative	
Notes	

Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			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		
			Low
DOMAIN 4: Flow and Timing			

Chen 2007 (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		

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 Oct 2000 and March 2001 were included
Patient characteristics and setting	Age: 58.3 (range 7-76) 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 Prevalance: 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

Cheng 2003a (Continued)

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 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			
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		
			Unclear
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		

Cheng 2003a (Continued)

Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?				
				Low
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			
				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			
Were all patients included in the analysis?	Yes			

Cheng 2003b

Study characteristics	
Patient sampling	Second entry created for Cheng 2003, the 60 patients were split into two 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	

Cheng 2003b (Continued)

Target condition and reference standard(s)			
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?			
Did the study avoid inappropriate exclusions?			
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?			
Were there any conflicts of interest? (if no information given at all then judge as unclear)			
Was calibration of examiners undertaken and results reported?			
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?			

Cheng 2003b (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?			
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?			
Did all patients receive the same reference standard?			
Were all patients included in the analysis?			

Delavarian 2010

Study characteristics	
Patient sampling	<p>Method of patient selection: Quote “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 Otorhinolaryngology Department of QAEM, IMAM, REZA and OMID hospitals, Mishhad, Iran.”</p> <p>Quote “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.”</p>
Patient characteristics and setting	<p>Age: 54.00 ± 17.38</p> <p>Sex: 13 men 12 women</p> <p>SES: Not reported</p> <p>Ethnicity: Not reported</p> <p>Stated Risk Factors: unclear</p> <p>Number of patients/lesions: 25/26</p> <p>Lesion site: Oral cavity</p> <p>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”</p> <p>Country: Iran</p> <p>Type of facility: Secondary</p> <p>Prevalance: (dysplasia / malignancy) 8/26</p> <p>Exclusions: history of any treatment for the lesion, systemic contraindication for scalpel biopsy</p>

Index tests	<p>Category: Cytology - OralCDx (lab processing not performed at OralCDx lab)</p> <p>Description: Quote "After determination of site biopsy, under local anaesthesia, needed for scalpel biopsy, the Oral CDx brush was placed in the selected area and turned 5 to 10 times until appearing pinpoint bleeding-upon manufacture's recommendation."</p> <p>Positivity threshold: Quote "The pathological findings were categorized as three groups: 1) Positive: dysplastic epithelial changes 2) Negative: absence of any evidence suggesting dysplasia 3) Inadequate sampling"</p> <p>Sequence of tests: Index followed by reference</p> <p>Training or calibration of clinicians: Unclear. quote "They were examined by a pathologist informed about clinical diagnosis"</p> <p>Blinding of examiners: Index completed prior to reference, quote "blind to the histopathological results"</p> <p>Multiple tests: No</p> <p>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"</p> <p>Conflict of interests: University support acknowledged.</p>
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: Scalpel biopsy method unclear.</p> <p>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)."</p> <p>Sequence of tests: index followed by reference</p> <p>Training or calibration of pathologists: unclear</p> <p>Blinding of examiners: Quote "The histopathologic preparations were observed by the same pathologist blind to the cytopathical study and informed about clinical diagnosis"</p> <p>Multiple tests: not applicable</p> <p>Method of site selection: Quote "The scalpel biopsy was done immediately in the site of pin-point bleeding."</p> <p>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)</p>
Flow and timing	<p>Patients receiving index test but not reference test: 0</p> <p>Patients receiving reference test but not index test: 0</p> <p>Time interval: Not discussed</p> <p>Patients receiving both index and reference test but excluded from analysis: 0</p>
Comparative	
Notes	<p>Assumed figures in Table 2 should be reversed: "Gold Standard" Normal and Disease should be reversed. This has been done for data entry</p>
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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
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		
			Low
DOMAIN 4: Flow and Timing			

Delavarian 2010 (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?	Yes		

Du 2007

Study characteristics	
Patient sampling	<p>Method of patient selection: 132 patients from School and Hospital of Stomatology, Wuhan University, between July 2002 and November 2003</p> <p>Inclusion criteria for patients:</p> <ul style="list-style-type: none"> (i) superficial ulceration suspicious of malignancy (ii) oral leukoplakia (iii) oral lichen planus (iv) oral leukokeratosis <p>Excluded:</p> <ul style="list-style-type: none"> (i) benign oral lesion (ii) lesion without histological result after clinical diagnosis
Patient characteristics and setting	<p>Age: Mean 49.4 SD 12.7</p> <p>Sex: 67 female, 61 male</p> <p>SES: not specified</p> <p>Ethnicity: not specified</p> <p>Stated Risk Factors: not discussed</p> <p>Number of patients/lesions: 128/128</p> <p>Lesion site: Oral cavity</p> <p>Severity: All patients suspected of malignancy, leukoplakia, lichen planus, leukokeratosis</p> <p>Country: China</p> <p>Type of facility: Secondary</p> <p>Prevalance: 33/128</p>
Index tests	<p>Category: Vital staining - Rose Bengal</p> <p>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."</p> <p>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."</p>

	<p>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</p>		
Target condition and reference standard(s)	<p>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 one experienced oral pathologist used. Blinding of examiners: Blind to index test. Multiple tests: No. Method of site selection: Unclear but each patient had one lesion, assume biopsy taken from this lesion Target condition: Malignant or benign.</p>		
Flow and timing	<p>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</p>		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		

Du 2007 (Continued)

Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	No		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
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		
Low			
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Study characteristics	
Patient sampling	Method of patient selection: Quote "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)" quote "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) 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 = two thirds, 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: US Type of facility: secondary care Prevalance: 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 30-60 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 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 blind 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"

Epstein 2008 (Continued)

	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 won't have significant effect on results		
Comparative			
Notes	Concerns regarding this study analysis for cancer only, ignoring PMDs. 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			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Yes		
Was calibration of examiners undertaken and results reported?	No		

Epstein 2008 (Continued)

Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?	No			
Low				
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			
Low				
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			

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 Vizi-Lite"
Patient characteristics and setting	Setting: Oral medicine specialist, secondary care. Age: Male 56.81 (2.2), female 58.7 (2.47) Sex: Male 47%, Female 53% SES: not discussed Ethnicity: not discussed Stated Risk Factors: 21 smokers (38%)

	<p>Number of patients/lesions:55/80, only reported on primary lesions = 55 Lesion site: Aleolar 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 Prevalance: 10/55</p>
<p>Index tests</p>	<p>Category: Light based - Vizilite Description: Quote "After a 60s 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: Intra-oral examinations Conflict of interests:none</p>
<p>Target condition and reference standard(s)</p>	<p>Category:Biopsy with histopathological assessment Description: Quote "an incisional scalpel biopsy was performed under local anaesthesia to obtain a definitive histopathological diagnosis" 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</p>
<p>Flow and timing</p>	<p>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</p>

Comparative			
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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge	Yes		

Farah 2007 (Continued)

of the results of the index tests?			
			Low
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		

Farah 2012

Study characteristics	
Patient sampling	Method of patient selection: Quote: "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) 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 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" Prevalance: 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

	<p>were deemed negative for loss of autofluorescence only if complete blanching was achieved.”</p> <p>Sequence of tests: Index test followed by reference standard</p> <p>Training or calibration of clinicians: Calibration of VELScope findings are reported, but no reporting of training</p> <p>Quote: “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.”</p> <p>Quote: “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.”</p> <p>Blinding of examiners: Index test completed prior to reference standard. Nothing specifically reported</p> <p>Multiple tests: No</p> <p>Method of site selection: visual examination</p> <p>Conflict of interests: Not specifically reported.</p>		
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessments</p> <p>Description: Quote: “A scalpel tissue biopsy was taken.”... “Biopsy specimens were fixed in formalin, blocked in paraffin, stained with hematoxylineosin, and assessed by routine histopathology ”</p> <p>Positivity threshold: Taken from table 4, results of histology worked back to clinical definitions</p> <p>Sequence of tests: Index test followed by reference standard</p> <p>Training or calibration of pathologists: Quote: “There was a 96.6% interobserver agreement for histopathological interpretation between the 2 pathologists.”</p> <p>Blinding of examiners: Quote: “Both examiners were blinded to the clinical findings”</p> <p>Multiple tests: No</p> <p>Method of site selection: Quote: “Loss of autofluorescence was used to determine the best site for biopsy”</p> <p>Target condition: Dysplasia and OSCC</p>		
Flow and timing	<p>Patients receiving index test but not reference test: 0</p> <p>Patients receiving reference test but not index test: 0</p> <p>Time interval: Not reported.</p> <p>Patients receiving both index and reference test but excluded from analysis:0</p>		
Comparative			
Notes			
Methodological quality			
Item	Authors’ judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		

		Low	
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Yes		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
		Low	
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		
		Low	
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		

Study characteristics	
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	<p>Age: mean 56.2 years Sex: 13 men, 22 women 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 Prevalance: 15/43</p>
Index tests	<p>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: 20s pre-rinse with 30ml of 1% acetic acid; 20s water rinse; 20s rinse/gargle with 10 ml of the 1% tolonium chloride solution; 20s post-rinse with 30ml 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 one 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</p> <p>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” Positivity threshold: Quote “The brush cytology results were classified as malignant, atypical (suspicious), benign tissues or inadequate sample” Sequence of tests: TB, 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</p>

	<p>of first index test used to inform second index test</p> <p>Method of site selection: not reported</p> <p>Conflict of interests: stated no conflict and funded by University</p>
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: Quote "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."</p> <p>Positivity threshold: Quote "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'"</p> <p>Sequence of tests: Index then reference</p> <p>Training or calibration of pathologists: Quote "Surgical biopsies were performed by an experienced oral and maxillofacial surgeon."</p> <p>Blinding of examiners: Not reported.</p> <p>Multiple tests: One reference standard.</p> <p>Method of site selection: Quote "All areas retaining Tblue were biopsied; in sites with no retention of staining, clinical judgment guided the biopsy procedure."</p> <p>Target condition: 'Pathologic interpretation was based on established criteria (WHO) and classified as squamous cell carcinoma, epithelial dysplasia, hyperkeratosis, lichen planus, and other benign lesions.'</p>
Flow and timing	<p>Patients receiving index tests but not reference test: 0</p> <p>Patients receiving reference test but not index tests: 0</p> <p>Time interval: not more than two weeks between the three methods of investigation</p> <p>Patients receiving both index and reference test but excluded from analysis: one for brush biopsy considered a false positive because of inadequate cell sampling</p>
Comparative	
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			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
			Low
DOMAIN 2: Index Test Test group A			

Guneri 2011 (Continued)

If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			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		
			Low
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		

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	<p>Age: (benign and premalignant) 38, range 19-75 years; (Squamous Cell Carcinoma) 52, range 35-74 years</p> <p>Sex: (benign and premalignant) 42 men 22 women; (SCC) 24 men 8 women</p> <p>SES: not reported</p> <p>Ethnicity: not reported (India)</p> <p>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</p> <p>Number of patients/lesions: 96</p> <p>Lesion site: Oral cavity</p> <p>Severity: Quote "suspicious pre-malignant or malignant lesions of the oral cavity irrespective of site, stage and sex were selected"</p> <p>Country: India</p> <p>Type of facility: Secondary</p> <p>Prevalance: (pre-malignant and malignant) 46/96</p>
Index tests	<p>Index Test 1: Toluidine Blue</p> <p>Category: Vital staining</p> <p>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."</p> <p>Positivity threshold: Participants' results classed as positive or negative but no thresholds or inadequate/equivocal results reported</p> <p>Sequence of tests: Staining, brush biopsy, followed by reference standard</p> <p>Training or calibration of clinicians: not reported</p> <p>Blinding of examiners: Index completed prior to reference.</p> <p>Multiple tests: yes. Combined test results used (Table 3) results of first index test used to inform second index test</p> <p>Method of site selection: not reported</p> <p>Conflict of interests: not reported</p> <p>Index Test 2: Brush Cytology</p> <p>Category: Cytology</p> <p>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."</p> <p>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 pattern and distribution, and discrepancy in maturation." Participants' results classed as positive or negative but no thresholds or inadequate / equivocal results reported</p> <p>Sequence of tests: TB, brush biopsy, followed by reference standard</p> <p>Training or calibration of clinicians: not reported</p>

	<p>Blinding of examiners: Index completed prior to reference. Not reported whether results of TB 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.</p>		
Target condition and reference standard(s)	<p>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), pre-malignant (dysplasia) and squamous cell carcinoma</p>		
Flow and timing	<p>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</p>		
Comparative			
Notes	<p>Exclude brush biopsy as an independent index test (blue staining present) and only use results for toluidine blue staining and combined index test of staining and brush biopsy. Numbers for TP, FP, TN, FN have been calculated from raw data presented in table III sensitivity and specificity values. Results have been entered for both premalignant and malignant</p>		
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		
			Unclear
DOMAIN 2: Index Test Test group A			

If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
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		
			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		
Were all patients included in the analysis?	Yes		

Koch 2011a

Study characteristics	
Patient sampling	Method of patient selection: Quote "All patients attended the Maxillofacial Surgery Clinic at the Univeristy 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 Prevalance: 113/182
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: Two 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 two 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

Koch 2011a (Continued)

	Patients receiving both index and reference test but excluded from analysis: 4		
Comparative			
Notes	Low attrition		
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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Koch 2011a (Continued)

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

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”
Patient characteristics and setting	Age: Mean 61.7 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 palatogloss. 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 nm) in normal mucosa.” Documented by digital reflex photography

	<p>Positivity threshold: SCC, and dysplasia [identified] depending on two different autofluorescence features:</p> <p>(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</p> <p>(2) The presence of red mucosal autofluorescence was evaluated as a separate indicator for dysplasia or SCC (positive)</p> <p>Sequence of tests: Index then reference</p> <p>Training or calibration of clinicians: Not clearly discussed</p> <p>Blinding of examiners: One experienced examiner performed the visual examination, two examiners examined the photographs, they were blinded and independent</p> <p>Multiple tests: Visual then light, class as one index test only, as an adjunct</p> <p>Method of site selection: Not discussed</p> <p>Conflict of interests: None</p>		
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>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.”</p> <p>Positivity threshold: Mucosal hyperkeratosis, Lichen planus, SCC, Inflammation, Dysplasia, Healthy mucosa</p> <p>Sequence of tests: Index then Reference</p> <p>Training or calibration of pathologists: performed by one experienced examiner</p> <p>Blinding of examiners: not described</p> <p>Multiple tests: no</p> <p>Method of site selection: biopsy of photographed lesion</p> <p>Target condition: SCC and dysplasia</p>		
Flow and timing	<p>Patients receiving index test but not reference test: 0</p> <p>Patients receiving reference test but not index test: 0</p> <p>Time interval: Biopsy taken immediately after index test.</p> <p>Patients receiving both index and reference test but excluded from analysis: 0</p>		
Comparative			
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		

				Low
DOMAIN 2: Index Test Test group A				
If a threshold was used, was it pre-specified?	Yes			
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No			
Was calibration of examiners undertaken and results reported?	Unclear			
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?				
				Low
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			
				Low
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			

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	<p>Setting: Not clearly stated</p> <p>Age: 57.7, range 37-74</p> <p>Sex: not discussed</p> <p>SES: not discussed:</p> <p>Ethnicity: not discussed</p> <p>Stated Risk Factors: Quote "All patients had a history of smoking and of drinking alcohol"</p> <p>Number of patients/lesions: 58/160</p> <p>Lesion site: oral cavity</p> <p>Severity: not discussed</p> <p>Country: Germany</p> <p>Type of facility: not discussed</p> <p>Prevalance: 100/160</p>
Index tests	<p>Category: Light based</p> <p>Name of test(s): Imaging 5-Aminolevulinic Acid-Induced Protoporphyrin IX Fluorescence</p> <p>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)."</p> <p>Positivity threshold: strong, macroscopically visible red fluorescence (F++), weak (F+) or negative (F-)</p> <p>Sequence of tests: index then reference</p> <p>Training or calibration of clinicians: unclear</p> <p>Blinding of examiners: unclear</p> <p>Multiple tests: unclear</p> <p>Method of site selection: light-based detection</p> <p>Conflict of interests: not discussed</p>
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: Quote "Biopsy specimens were taken from tumor, tumor boundaries, and normal tissue under fluorescence illumination."</p> <p>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"</p> <p>Sequence of tests: index then reference</p> <p>Training or calibration of pathologists: not discussed</p> <p>Blinding of examiners: not discussed</p> <p>Multiple tests: no</p> <p>Method of site selection: specimens taken from identified sites</p> <p>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."</p>

Leunig 2000 (Continued)

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			
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		
			Unclear
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Leunig 2000 (Continued)

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

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 lesions 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: US Type of facility: Secondary - Veteran Administration Medical Centre so concern over wider applicability Prevalance: 105/235
Index tests	Category: Vital staining - Toluidine Blue Description: Photographed before and after direct application of stain, 2x mouth rinse with water (20 secs each), 1 rinse with 1% acetic acid solution (20 secs), area dried with gauze, '1% toluidine

Mashberg 1980 (Continued)

	<p>blue solution containing acetic acid and alcohol' applied with cotton swab, rinse with acetic acid solution (1 min), rinse with water</p> <p>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."</p> <p>Sequence of tests: TB then biopsy</p> <p>Training or calibration of clinicians: not discussed, possibly only one examiner</p> <p>Blinding of examiners: not explicitly stated but implied as biopsy results are returned for comparison with TB results</p> <p>Multiple tests: No</p> <p>Method of site selection: Not described, assumed sites of lesions after examination</p> <p>Conflict of interests: not reported</p>
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: 'A biopsy of the stained areas was performed' or 'a biopsy of the most suspicious area as defined by our erythroplastic criteria was performed'</p> <p>Positivity threshold: Invasive carcinoma or Carcinoma in situ</p> <p>Sequence of tests: index then reference</p> <p>Training or calibration of pathologists: not discussed</p> <p>Blinding of examiners: not discussed</p> <p>Multiple tests: no</p> <p>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"</p> <p>Target condition: Invasive carcinoma, Carcinoma in situ, Atypia, Benign</p>
Flow and timing	<p>Patients receiving index test but not reference test: 0</p> <p>Patients receiving reference test but not index test: 0</p> <p>Time interval: description implied biopsy taken immediately after rinsing</p> <p>Patients receiving both index and reference test but excluded from analysis: 0</p>
Comparative	
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		
			Unclear
DOMAIN 2: Index Test Test group A			

Mashberg 1980 (Continued)

If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
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		
			Low
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		

Study characteristics	
Patient sampling	Method of patient selection: Quote "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 26-87.2) 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 Quote: "Location of primary and satellite lesions is listed in Table 2, with the majority of primary lesions occurring on the tongue and buccal mucosa." Quote "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: Queensland, 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-head-light, and handheld re-usable LED diffused white-light guide to illuminate irregular cells Quote: "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" Positivity threshold: Quote: "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."

McIntosh 2009 (Continued)

	<p>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.</p>		
Flow and timing	<p>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</p>		
Comparative			
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			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index	No		

McIntosh 2009 (Continued)

test?			
			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		
			Low
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		

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 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: ‘Only lesions with an abnormal epithelial surface including erythroplakia, leukoplakia without dysplasia and oral submucous fibrosis were included’

Mehrotra 2008 (Continued)

	Country: India Type of facility: Medical College Prevalance: 34/79
Index tests	Category: Cytology - baby toothbrush Description: A hard nylon baby tooth brush, using moderate pressure, was repeatedly brushed in one 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
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	
Notes	
Methodological quality	
Item	Authors' judgement Risk of bias Applicability concerns

Mehrotra 2008 (Continued)

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			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		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Mehrotra 2008 (Continued)

Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	No			

Mehrotra 2010

Study characteristics	
Patient sampling	<p>Method of patient selection: Patients selected for study after detection of a clinically innocuous lesion (Class II) during routine dental care</p> <p>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.”</p>
Patient characteristics and setting	<p>Age: mean age 40 years</p> <p>Sex: 28 female, 230 male</p> <p>SES: Not stated</p> <p>Ethnicity: Not stated</p> <p>Stated Risk Factors: Tobacco usage 70.5%</p> <p>Number of patients/lesions: 258/258 split into two groups of 102 ViziLite+, 156 VELscope</p> <p>Lesion site: Buccal mucosa, retromolar trigone, tongue, alveolar mucosa, gingiva, floor of mouth, palate</p> <p>Severity: Identified as Class II prior to biopsy, so patients classified as Class I (quote “suspicious enough to warrant a biopsy”) were excluded</p> <p>Country: India</p> <p>Type of facility: Quote: “outpatient department of the government-run District Hospital ”</p> <p>Prevalance: 12/156 & 4/102, combined 116/258</p>
Index tests	<p>Index test 1 - VELScope</p> <p>Category: Light-based</p> <p>Description: Quote: “The clinicians performed the examinations with the VELscope and ViziLite devices according to the manufacturers’ instructions”</p> <p>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.”</p> <p>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</p> <p>Training or calibration of clinicians: Calibrated by an experienced professional from the manufacturer, no results reported</p> <p>Blinding of examiners: Index test examiners blind to reference test results.</p> <p>Multiple tests: Yes; conducted independently during the same session, by the same examiners</p> <p>Method of site selection: Principle area of morphology decided by consensus after visual examination</p> <p>Conflict of interests: Quote: “None of the authors reported any disclosures.”</p>

	<p>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 blind 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."</p>		
Target condition and reference standard(s)	<p>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. Training or calibration of pathologists: Not reported, although results interpreted by two 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 one of two); biopsy Method of site selection: Not specifically reported. Target condition: Quote: "potentially dysplastic and cancerous oral lesions"</p>		
Flow and timing	<p>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</p>		
Comparative			
Notes	<p>Velscope and Vizilite plus tests reported independently so data extracted for both</p>		
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		

				Low
DOMAIN 2: Index Test Test group A				
If a threshold was used, was it pre-specified?	Yes			
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No			
Was calibration of examiners undertaken and results reported?	No			
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?	Yes			
				Low
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			
				Low
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			

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%). 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 Prevalance: 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: Three 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’.

Mehrotra 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: reference test biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: 6 (inadequate brush biopsy sample)		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
			Unclear
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
DOMAIN 3: Reference Standard			

Mehrotra 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?	Yes		
Low			
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		

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, 23-80 years). Sex: 9 females and 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	<p>Category: Vital staining plus adjunct - Vizilite Plus</p> <p>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 assess toluidine blue retention</p> <p>Positivity threshold: Quote “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 coloration equates to negative, incomplete or positive</p> <p>Sequence of tests: Index then reference</p> <p>Training or calibration of clinicians: not discussed</p> <p>Blinding of examiners: not discussed</p> <p>Multiple tests: one combined test. Data taken from combined results in Table 4</p> <p>Method of site selection: visual and chemiluminescent examination</p> <p>Conflict of interests: none stated</p>		
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: Quote “incisional biopsy under local anaesthesia, and sample tissues on which biopsy was performed were submitted for histopathologic assessment”</p> <p>Positivity threshold: Positive result: Dysplasia, SCC; Negative: Benign keratosis, Lichen planus, Normal oral mucosa</p> <p>Sequence of tests: index then reference</p> <p>Training or calibration of pathologists: not described</p> <p>Blinding of examiners: not described</p> <p>Multiple tests: no</p> <p>Method of site selection: visual examination of index test</p> <p>Target condition: dysplasia or squamous cell carcinoma</p>		
Flow and timing	<p>Patients receiving index test but not reference test: 0</p> <p>Patients receiving reference test but not index test: 0</p> <p>Time interval: not reported</p> <p>Patients receiving both index and reference test but excluded from analysis: 0</p>		
Comparative			
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		

					Low
DOMAIN 2: Index Test Test group A					
If a threshold was used, was it pre-specified?	Unclear				
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear				
Was calibration of examiners undertaken and results reported?	Unclear				
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?	No				
					Low
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				
					Low
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				

Study characteristics	
Patient sampling	<p>Method of patient selection:</p> <p>Quote “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.”</p> <p>Quote: “A provisional diagnosis of leukoplakia, speckled leukoplakia, erosive lichen planus (pre-malignant lesions) and oral malignancies were made on basis of clinical examination.”</p>
Patient characteristics and setting	<p>Age: not reported</p> <p>Sex: not reported</p> <p>SES: not reported</p> <p>Ethnicity: not reported (India)</p> <p>Risk Factors: not stated</p> <p>Number of patients/lesions: 30/30 with clinically suspicious premalignant lesions and 30/30 with clinically suspicious malignant lesions</p> <p>Lesion site: oral cavity</p> <p>Severity: premalignant lesions (degree of dysplasia), malignant lesions (degree of differentiation)</p> <p>Country: India</p> <p>Type of facility: Department of Oral Medicine and Radiology, Bapuji Dental College and Hospital, Davangere, Karnataka</p> <p>Prevalence: 55/60</p>
Index tests	<p>Category: Vital staining - Toluidine-Blue combined with Lugol’s iodine</p> <p>Description: Quote “The subjects comfortably seated in the dental chair were examined following the methods described by Kerr et al.”</p> <p>Positivity threshold: Either or both of the tests staining positive.</p> <p>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”</p> <p>Lugol: 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.’</p> <p>Sequence of tests: TB Staining followed by Lugol’s iodine followed by reference standard</p> <p>Training or calibration of clinicians: No training reported</p> <p>Blinding of examiners: Index test completed before reference standard.</p> <p>Multiple tests: No but a combination of 2 tests.</p> <p>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”</p> <p>Conflict of interests: Not stated</p>
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: Quote: “Histopathological grading for premalignant lesions was performed as per the pathologic features suggested by Axell et al.”</p> <p>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).”</p> <p>Sequence of tests: Index test followed by reference standard</p>

	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		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
			Unclear
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			

				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			
				Low
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			

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) 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 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, one 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			
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		
			Unclear
DOMAIN 2: Index Test Test group A			

Navone 2004 (Continued)

If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
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		
			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		

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 six 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 Prevalance: 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 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: Three 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	
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) 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		
			Unclear
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted with-			

Navone 2008 (Continued)

out knowledge of the first index test?				
				Low
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			
				Low
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			

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

	Prevalance: 28/171
Index tests	<p>Category: Cytology - Oral Advance</p> <p>Description: QC samples were collected using the Oral Advance kit, using a flexible cytology brush. Image analysis followed by visual cytomorphologic analysis</p> <p>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</p> <p>Sequence of tests: Index followed by reference.</p> <p>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</p> <p>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."</p> <p>Multiple tests: No</p> <p>Method of site selection: Quote: "multiple locations of the suspicious lesion 10 times in one direction."</p> <p>Conflict of interests: One of the authors is a general pathology consultant at Perceptronix Medical, the company offering the Oral Advance index test</p>
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: Surgical biopsy "Oral biopsies were then taken from the same lesion in a conventional surgical fashion."</p> <p>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"</p> <p>Sequence of tests: Index then reference.</p> <p>Training or calibration of pathologists: Not discussed.</p> <p>Blinding of examiners: Yes "Nonaffiliated 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."</p> <p>Multiple tests: No</p> <p>Method of site selection: Oral biopsy taken from the same site as exfoliative cytology</p> <p>Target condition: Severe dysplasias and carcinoma in situ.</p>
Flow and timing	<p>Patients receiving index test but not reference test: 0</p> <p>Patients receiving reference test but not index test: 0</p> <p>Time interval: Reference test biopsy immediately followed the index test quote "concurrent QC assessment"</p> <p>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)</p>
Comparative	
Notes	
Methodological quality	

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Yes		
Was calibration of examiners undertaken and results reported?	No		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
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		
			Low
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		

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 23-92 years Sex: 77 men 53 women SES: not reported Ethnicity: Japanese Stated Risk Factors: not reported Number of patients/legions: 130 Lesion site: Oral cavity Severity: unclear Country: Japan Type of facility: Secondary Prevalance: 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

Onizawa 1999 (Continued)

	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			
Notes	Data extractable for oral cancer as target condition		
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		
			Unclear
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Unclear

Onizawa 1999 (Continued)

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		
		Low	
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		

Onofre 2001

Study characteristics	
Patient sampling	Method of patient selection: Quote “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)” Quote “Not included in this study were 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” Prevalance: 13/50
Index tests	<p>Category: Vital staining - Toluidine blue</p> <p>Description: Quote: “All lesions were submitted to staining with an aqueous solution of 1% toluidine blue, ..” Followed recommendations of Mashberg (Mashberg 1980).</p> <p>Positivity threshold: Followed recommendations of Mashberg (Mashberg 1980, Mashberg, 1983)</p> <p>1) “inadequate cell count” 2) “negative” 3) “atypical epithelial cells” 4) “positive for dysplasia or OSCC”</p> <p>Atypical and positive results recorded as positive; inadequate results excluded</p> <p>Sequence of tests: Index test followed by reference standard.</p> <p>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.”</p> <p>Blinding of examiners: Not explicitly reported but index test was performed prior to reference standard</p> <p>Multiple tests: No</p> <p>Method of site selection: Quote: “sites were selected on the basis of the clinical appearance of the lesion and the staining result”</p> <p>Conflict of interests: Quote “We are indebted to the Mario A.S. Paino Laboratory of Clinical Pathology.”</p>
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: biopsy and histologic analysis</p> <p>Positivity threshold: any grade of dysplasia (in OL and OLP) and OSCC</p> <p>Sequence of tests: Index test preceded reference standard.</p> <p>Training or calibration of pathologists: Quote: “Previously calibrated pathologist.”</p> <p>Blinding of examiners: Quote: ”pathologist was not informed of the staining result.”</p> <p>Multiple tests: no</p> <p>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.”</p> <p>Target condition: Squamous cell carcinomas, Epithelial dysplasia, Keratosis, Lichen Planus</p>
Flow and timing	<p>Patients receiving index test but not reference test: 0</p> <p>Patients receiving reference test but not index test: 0</p> <p>Time interval: not reported</p> <p>Patients receiving both index and reference test but excluded from analysis: 0</p>
Comparative	
Notes	Clinical interpretation of cutoff required.
Methodological quality	
Item	Authors’ judgement Risk of bias Applicability concerns
DOMAIN 1: Patient Selection	

Onofre 2001 (Continued)

Was a consecutive or random sample of patients enrolled?	No			
Did the study avoid inappropriate exclusions?	No			
				Low
DOMAIN 2: Index Test Test group A				
If a threshold was used, was it pre-specified?	Yes			
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear			
Was calibration of examiners undertaken and results reported?	Yes			
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?				
				Low
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			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Unclear			

Onofre 2001 (Continued)

Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

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	<p>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</p> <p>Index Test 2 - Cytobrush Category: Cytology (following toluidine blue) Description: Quote "Visually identified lesions were then scraped using a cytobrush" "smears were... .fixed in 95% alcohol or air dried and stained with hermatoxylin-eosin, Pap and PAS stains" 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</p>

	<p>Blinding of examiners: All specimens were examined by two pathologists in a double blind fashion, any discrepancies were resolved by a third opinion Multiple tests: Yes Method of site selection: Visual and staining Conflict of interests: Not discussed</p>
Target condition and reference standard(s)	<p>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: Blind 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</p>
Flow and timing	<p>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</p>
Comparative	
Notes	<p>Study is a direct comparison of two 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</p>

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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		

Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?	No		
Low			
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		
Low			
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		

Remmerbach 2009

Study characteristics	
Patient sampling	Method of patient selection: 47 patients referred for examination and treatment of oral lesions were examined by an expert from Dept of Oral and Maxillofacial Surgery, University of Leipzig. Potential selection of patients quote "After a thorough intraoral examination"
Patient characteristics and setting	Age: mean age 61, range 27-84 years, SD 12years Sex: not discussed SES: not reported Ethnicity: not reported Stated Risk Factors: not reported Number of patients/legions: 47 Lesion site: Oral cavity Severity: referred for treatment of oral lesion Country: Germany Type of facility: Secondary Prevalance: 20/47
Index tests	Category: Cytology Description: After an examination 4 smears were taken with a brush-based cell collector, rolled on a glass slide and fixed with propanol. A series of different stains were applied: Papanicolaou and Feulgen Staining, DNA Measurements performed, then AgNOR Analysis. All tracked with Multimodal Cell Analysis Positivity threshold: '1) "insufficient" for specimens without any or with exclusively autolytic cells; 2) "tumor cell negative" for unsuspecting, reactive, or inflammatory cellular images; 3) "atypical" in cases with atypical cellular changes (e.g., with mild or moderate dysplasia); 4) "suspicious for tumor cells" if only sparse abnormal or severe dysplastic squamous cells were observed or if the diagnostic criteria for malignancy were only vague; and 5) "tumor cell positive" for smears that contained unequivocal malignant cells' (1) & (2) classified as negative (3) - (5) positive Sequence of tests: Index then reference test Training or calibration of clinicians: Not discussed, patients examined by one person only who is described as experienced Blinding of examiners: Only one examiner, not discussed whether blind of histology Multiple tests: Yes: Cytology, DNA, AgNOR Method of site selection: According to examination Conflict of interests: No conflict 'project is supported by the Viktor and Mirka Pollak Fund for Biomedical Engineering and the Innovation Fund of the University Hospital Leipzig'
Target condition and reference standard(s)	Category: Biopsy with histopathologic assessment. Description: Scalpel biopsy-based histology, which was validated by clinical follow-up Positivity threshold: Squamous cell carcinomas considered positive. Negative results - leukoplakias, lichen planus Sequence of tests: Index then reference test. Training or calibration of pathologists: Not discussed Blinding of examiners: Not discussed Multiple tests: No

Remmerbach 2009 (Continued)

	Method of site selection: Not described Target condition: Squamous cell carcinomas		
Flow and timing	Patients receiving index test but not reference test: none Patients receiving reference test but not index test: none Time interval: assumed reference test biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: none		
Comparative			
Notes	Assuming independence between the three index tests.		
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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
DOMAIN 3: Reference Standard			

Remmerbach 2009 (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		
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		
Were all patients included in the analysis?	Yes		

Scheer 2011

Study characteristics

Patient sampling	Method of patient selection: Quote "Oral and VELscope examinations were performed on 64 patients referred to the Department of Oral and Craniomaxillo-facial Surgery to rule out invasive squamous cell carcinoma." "Patients with advanced squamous cell carcinomas were excluded" Twenty patients with previous history raise concern that examiners are already aware of patients diagnosis
Patient characteristics and setting	Age: average age of 59.8 years Sex: 25 female and 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 Craniomaxillo-facial Surgery Prevalence: 12/64

Index tests	<p>Category: Light based - VELscope Description: Quote "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</p>		
Target condition and reference standard(s)	<p>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 under-went 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.</p>		
Flow and timing	<p>Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test:0 Time interval: Quote "All lesions under-went histopathological evaluation by a trained pathologist after examination with the VELscope" Patients receiving both index and reference test but excluded from analysis: 0</p>		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Did the study avoid inappropriate exclusions?	No			
				Low
DOMAIN 2: Index Test Test group A				
If a threshold was used, was it pre-specified?	Yes			
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No			
Was calibration of examiners undertaken and results reported?	Unclear			
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?				
				Low
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			
				Low
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			

Scheifele 2004**Study characteristics**

Patient sampling	<p>Method of patient selection: Quote: "80 consecutive patients between July 2002 and September 2003". Inclusion criteria: "(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."</p> <p>Only those that received a scalpel biopsy were included in the study, not full spectrum of disease leading to potential sampling bias</p>
Patient characteristics and setting	<p>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: Oral CDx 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" Prevalance: Dysplasia or carcinoma 26/96 (lesion level)</p>
Index tests	<p>Category: Cytology - Oral CDx 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) Quote "1) "inadequate cell count", 2) "negative", 3) "atypical epithelial cells", 4) "positive for dysplasia or OSCC" Quote: "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" 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."</p>

Scheifele 2004 (Continued)

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: One month Patients receiving both index and reference test but excluded from analysis: 7/103 Quote: “Overall, there were seven (6.8%) inadequate results.”		
Comparative			
Notes			
Methodological quality			
Item	Authors’ judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Yes		
Was calibration of examiners undertaken and results reported?	Unclear		

Scheifele 2004 (Continued)

Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?				
				Low
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			
				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			

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 range 18-83 Sex: male 443 (47%), female 502 (53%) SES: not reported

	<p>Ethnicity: not reported Stated Risk Factors: Cigarette use, Other Tobacco, Alcohol. Number of patients/regions: 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 componen" 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 Prevalance: 102/298 Note: although selecting those with a 'suspicious lesion' this is still an acceptable population</p>
<p>Index tests</p>	<p>Category: Cytology - Oral CDx Description: Quote "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, N.Y." 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</p>
<p>Target condition and reference standard(s)</p>	<p>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 reports 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</p>
<p>Flow and timing</p>	<p>Patients receiving index test but not reference test: Class 1 cases 0, but class 2 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</p>

Comparative			
Notes	All class 1 received both index and reference test, class 2 received index then reference - if positive to index. Can only look at results of class 1 In data, classified Positive & 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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Yes		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Sciubba 1999 (Continued)

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

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 Copmostela, 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 (12.14 SD) 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 Prevalance: 11/24
Index tests	Category: Brush cytology - OralCDx 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

	<p>of the sample.”</p> <p>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)</p> <p>Sequence of tests: Index followed by reference</p> <p>Training or calibration of clinicians: No details stated for training or calibration.</p> <p>Blinding of examiners: Not stated</p> <p>Multiple tests: No</p> <p>Method of site selection: Not stated</p> <p>Conflict of interests: Not stated</p>		
Target condition and reference standard(s)	<p>Category: Biopsy with histopathologic assessment</p> <p>Description: Surgical biopsy</p> <p>Positivity threshold: Negative - no epithelial alteration, positive - dysplasia</p> <p>Sequence of tests: Index followed by reference</p> <p>Training or calibration of pathologists: Not stated</p> <p>Blinding of examiners: Not stated</p> <p>Multiple tests: No</p> <p>Method of site selection: Not stated</p> <p>Target condition: Dysplasia</p>		
Flow and timing	<p>Patients receiving index test but not reference test: 0</p> <p>Patients receiving reference test but not index test: 0</p> <p>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.”</p> <p>Patients receiving both index and reference test but excluded from analysis: 0</p>		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		

Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear			
Was calibration of examiners undertaken and results reported?	No			
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?				
				Low
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			
				Low
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			

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- 81 Sex: Not reported 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: Quote: "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 Time interval: Assumed reference test biopsy immediately followed the index test Patients receiving both index and reference test but excluded from analysis: Two

Sharwani 2006a (Continued)

Comparative			
Notes	Badly reported. Used 1.2 threshold for results. 1.3 level results: TP 26, FN 5, TN 34, FP 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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			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	Unclear		

Sharwani 2006a (Continued)

of the results of the index tests?				
				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			
Were all patients included in the analysis?	No			

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- 67 Sex: 13 M, 12 F 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: Quote: "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."

	<p>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. Conflict of interests: Quote: "We are grateful to "The British Association of Oral and Maxillofacial Surgeons" who provided the funding for this study."</p>		
Target condition and reference standard(s)	<p>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. Quote: "ESS was used to examine the suspicious area of each of those patients prior to surgical biopsy." Quote: "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. Quote: "ESS was used to examine the suspicious area of each of those patients prior to surgical biopsy." Quote: "The 25 biopsies were taken from various oral sites, and examined histopathologically." Target condition: Malignant (including dysplasia) or carcinoma in situ</p>		
Flow and timing	<p>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: Ten</p>		
Comparative			
Notes	<p>Unclear why 10 histologically confirmed hyperkeratotic lesions were excluded from the analysis</p>		
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		

			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
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		
			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		

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” Country: US Type of facility: secondary oral medicine clinic Prevalance: 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: Categorized 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			
Notes	Dysplasia and carcinoma extracted as positive for histology		
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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Unclear		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
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		

Silverman 1984 (Continued)

				Low
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			

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 18-89 Sex: 51% women 49% men 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 Prevalance: 93/298
Index tests	Category: Cytology - Oral CDx Description: Brush biopsy with computer assisted method of analysis (Oral cdx) 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 Oral CDx, but involvement of Oral CDx labs 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		
Comparative			
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		
			Unclear
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Yes		

Svirsky 2002 (Continued)

Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
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		
			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		
Were all patients included in the analysis?	Unclear		

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” Frank malignancy, hypersensitivity to ingredients in the light examination and any systemic disease were all excluded

Patient characteristics and setting	<p>Age: 44.4 (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 Prevalence: 17/99</p>
Index tests	<p>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” CHBT = 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 one index test. Method of site selection: Conventional visual examination then chemiluminescence and staining Conflict of interests: No</p>
Target condition and reference standard(s)	<p>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</p>
Flow and timing	<p>Patients receiving index test but not reference test: 0 Patients receiving reference test but not index test: 0 Time interval: All conducted in one session Patients receiving both index and reference test but excluded from analysis: Unclear</p>
Comparative	
Notes	<p>Performed at lesion level</p>

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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
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		
			Low
DOMAIN 4: Flow and Timing			

Ujaoney 2012 (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		

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, range 31-75 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 sec, 1% acetic acid for 20 sec, rinse 5ml of 1% toluidine blue, rinse 1% acetic acid 20 sec, 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: Quote "A suitable incisional/punch biopsy was obtained on the basis of site retaining

	<p>the stain” “Biopsy was also obtained from those lesions which did not retain any stain but were clinically suggestive of a PMOL’s” 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 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.</p>		
Flow and timing	<p>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</p>		
Comparative			
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		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Yes		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	No		
Was calibration of examiners undertaken and results reported?	Unclear		

Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?				
				Low
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			
				Low
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			

Warnakulasuriya 1996**Study characteristics**

Patient sampling	Method of patient selection: Quote "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." Quote "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 SD 15 years Sex: 73 men; 29 women SES: not reported

	<p>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) Prevalance: 57/86 lesions</p>
<p>Index tests</p>	<p>Category: Vital staining - Toluidine blue 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: TB rinse - No site selection Conflict of interests: KASSW supported by Dunhill Medical Trust. Consumables in project funded by Zila Pharmaceuticals</p>
<p>Target condition and reference standard(s)</p>	<p>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-µm 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: '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.</p>
<p>Flow and timing</p>	<p>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</p>

	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 TB equivocal excluded from analysis		
Comparative			
Notes			
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test Test group A			
If a threshold was used, was it pre-specified?	Unclear		
Were there any conflicts of interest? (if no information given at all then judge as unclear)	Yes		
Was calibration of examiners undertaken and results reported?	Unclear		
Where multiple index tests were used were the results of the second index test interpreted without knowledge of the first index test?			
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Warnakulasuriya 1996 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes			
Low				
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Unclear			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	No			

ALA: aminolevulinic acid
 CIS: carcinoma in situ
 OLP: oral lichen planus
 COE: conventional oral examination
 H&E: haematoxylin and eosin
 OSCC: oral squamous cell carcinoma
 PMD: potentially malignant disorders
 PMELs: potentially malignant epithelial lesions
 TB: toluidine blue

Characteristics of excluded studies [ordered by study ID]

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 Hohlwel-Majert.

(Continued)

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'
Hohlweg-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 cannot be extracted
Jayaprakash 2009	Data compromised by the inclusion of normal/control lesions.
Kulapaditharon 1998	Insufficient data.
Lane 2012	Data unavailable for true positive, true negative, false positive, false positive. 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	Cannot obtain satisfactory data to calculate sensitivity and specificity
Maraki 2006	Cannot obtain satisfactory data to calculate sensitivity and specificity
Navone 2007	Presentation of 'inadequate' cells make the calculation of sensitivity and specificity impossible from the data provided
Navone 2009	Data not presented to allow interpretation of sensitivity/specificity
Nieman 2008	Difficulties in clarifying the date, no response from contact author
Poate 2004	60 of the 75 who has 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

(Continued)

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 level not patient
Remmerbach 2007	Reported on number of samples (smears) taken, not at patient level, same numbers of patients and lesions as 2004 study
Sandler 1964	Diagnosis of cytology was used to determine whether biopsy would be undertaken
Schwarz 2009	66 patients without suspected oral lesions were included.
Shklar 1970	DTA study on a highly selected group. Selection criteria was histological diagnosis of epidermoid carcinoma
Silverman 1992	Re-publication of 1984 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
Wang 2009	No data for sensitivity or specificity- also uses 40 normal controls

DTA = Diagnostic Test Accuracy

DATA

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

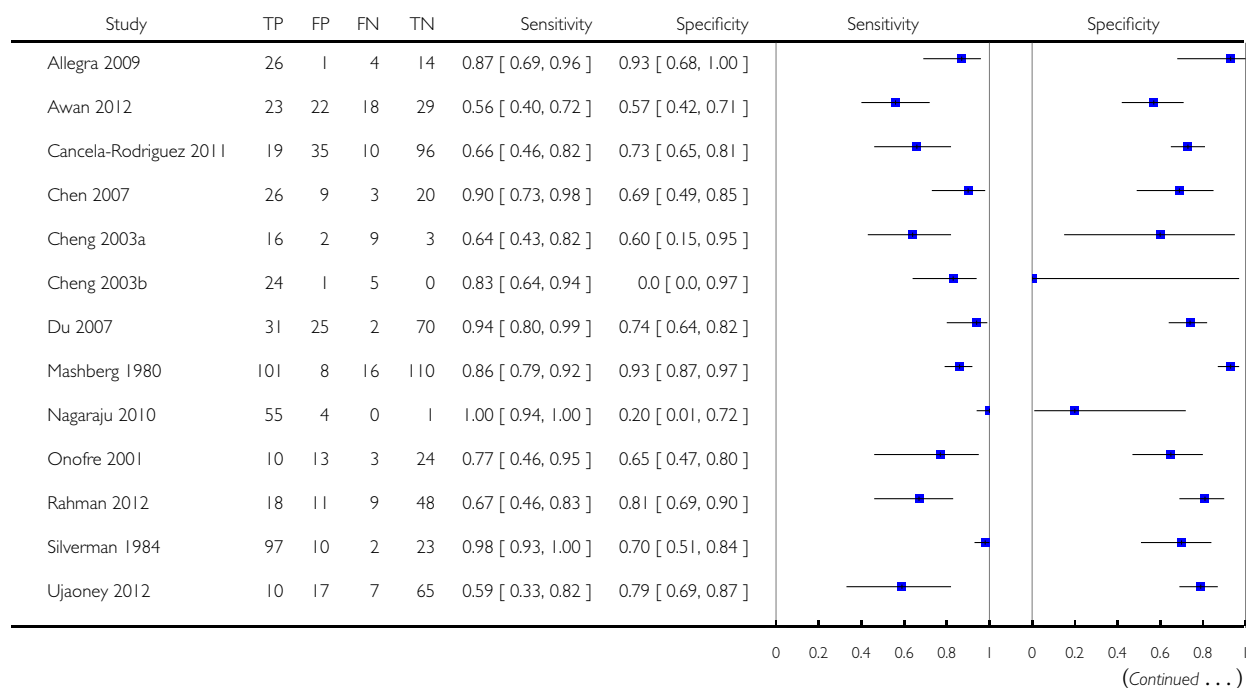
Tests. Data tables by test

Test	No. of studies	No. of participants
1 Vital staining	15	1338
2 Cytology	12	1575
3 Light-based	11	988
4 Vital staining plus adjunct (Light)	4	339
5 Vital staining plus adjunct (Cytology)	2	139

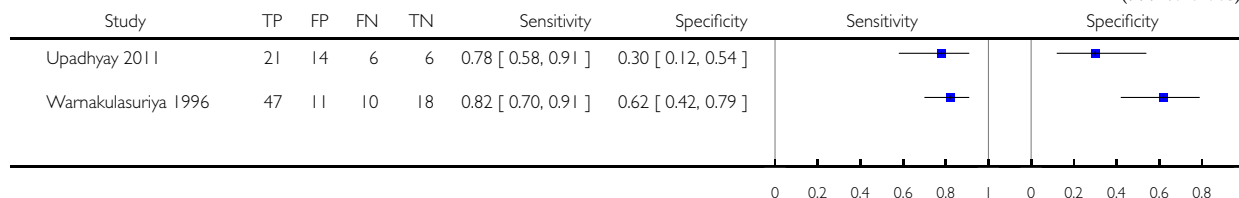
Test 1. Vital staining.

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

Test: 1 Vital staining



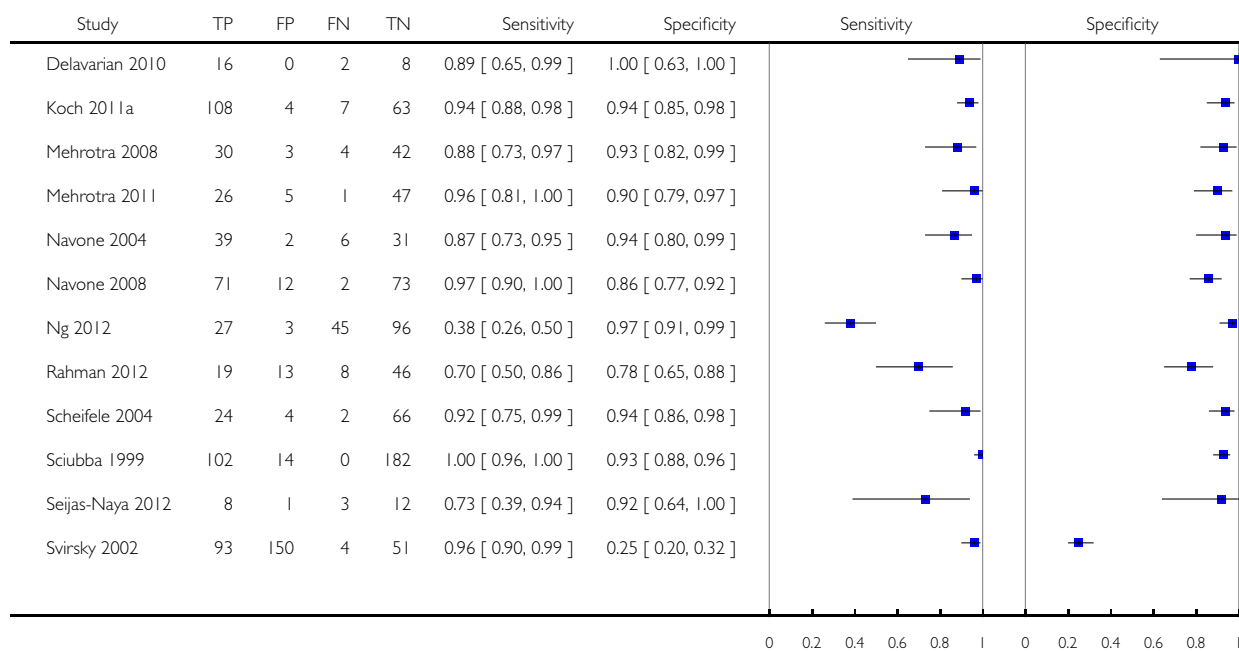
(... Continued)



Test 2. Cytology.

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

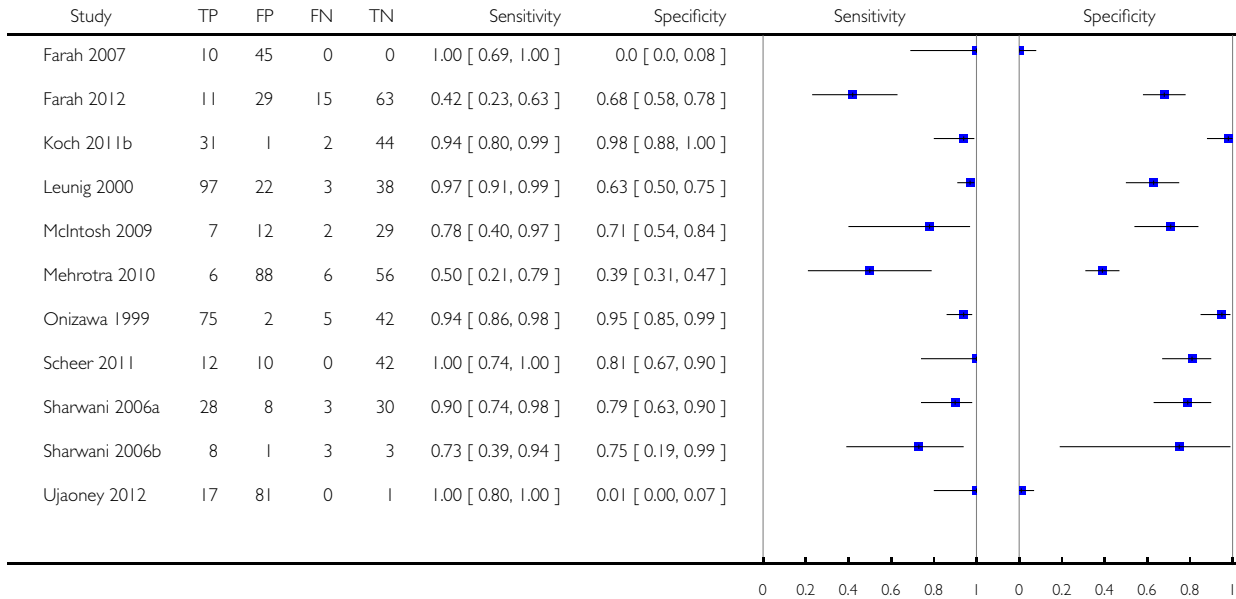
Test: 2 Cytology



Test 3. Light-based.

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

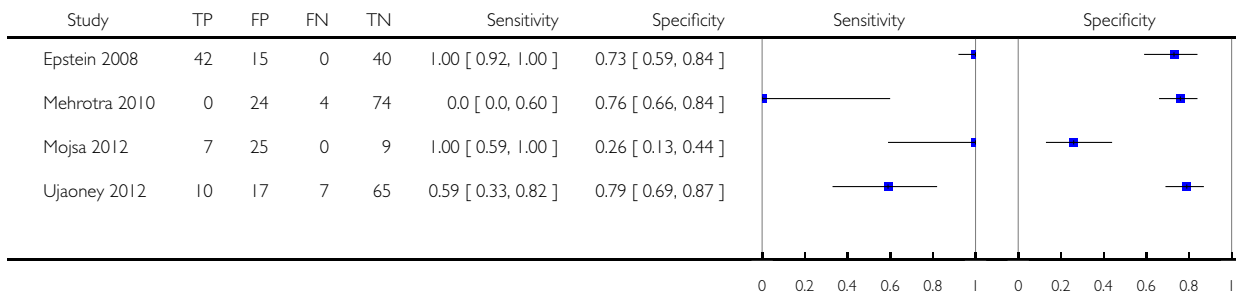
Test: 3 Light-based



Test 4. Vital staining plus adjunct (Light).

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

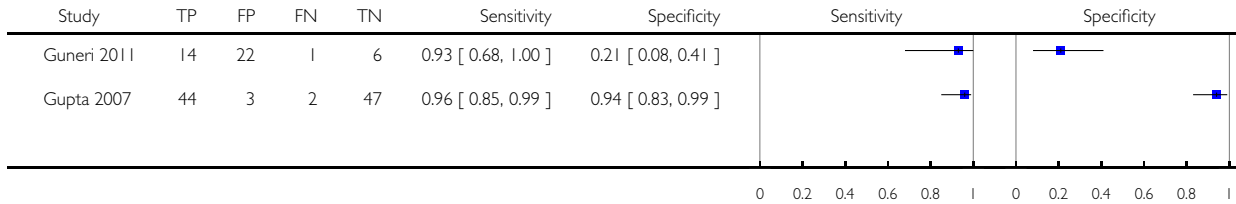
Test: 4 Vital staining plus adjunct (Light)



Test 5. Vital staining plus adjunct (Cytology).

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

Test: 5 Vital staining plus adjunct (Cytology)



ADDITIONAL TABLES

Table 1. Index tests for oral cancer and PMDs

Test	Characteristics	Classification of response	Other information
Conventional oral examination (COE)	A standard visual and tactile examination of the oral mucosa under normal (incandescent) light	The presence of an oral mucosal abnormality with a suspicion of malignancy or potential malignancy is classified as a positive test result; the presence of oral mucosal abnormality without a suspicion of malignancy or potential malignancy is classified as a negative test result	Traditionally used as a oral cancer screen rather than diagnosis, but its utility is debated (Lingen 2008). Advantages: quick and easy once trained, minimally invasive. Disadvantages: oral mucosal abnormalities are not necessarily clinically or biologically malignant; only a small percentage of leukoplakias are progressive or become malignant, COE cannot distinguish between those that are or are not; some precancerous lesions may exist within oral mucosa that appears clinically normal by COE alone (Lingen 2008).
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 PMD or malignancy (Leston 2010; Lingen 2008; Patton 2008). The procedure is as follows:	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 PMD for excision Disadvantages: benign inflammatory lesions are subject to

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

	<ul style="list-style-type: none"> • 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 		<p>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</p>
Brush cytology (e.g. OralCDx brush biopsy)	<p>Brush cytology refers to the microscopic assessment and interpretation of cell samples from PMD 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).</p>	<p>Following analysis, cytopathologists classify test results as positive, atypical or negative</p>	<p>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).</p> <p>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.</p>
Light-based detection (chemiluminescence e.g. ViziLite plus, tissue fluorescence imaging e.g. ViziLite, Microlux DL; VELscope, Identafi 3000; tissue fluorescence spectroscopy)	<p>Light-based systems to identify malignant and potentially malignant lesions and to highlight their presence through tissue reflectance (Leston 2010; Lingen 2008; Patton 2008) e.g. using Microlux DL, the procedure is as follows (Lingen 2008).</p> <ul style="list-style-type: none"> • Pre-rinse with acetic acid • Use blue-light source to visually assess the oral cavity. <p>ViziLite Plus also provides a toluidine chloride solution (toluidine blue) to aid in the marking of the lesion for biopsy</p>	<p>The result of the test is classed as negative if the appearance of the epithelium is lightly bluish white and positive if the appearance of the epithelium is distinctly white (acetowhite)</p>	<p>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</p> <p>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 re-</p>

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

	once the light source is removed		sults
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 PMD 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 PMDs

COE = conventional oral examination; PMD = 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	Describe the reference standard and how it was conducted and interpreted. Ideally, the biopsied tissue should be examined by more than one 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	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 scalpel 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

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

		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	the sequence of reference and index tests. Methods of site selection should be clearly documented	between the index test and reference standard should be short in the majority of cases. If the period elapsed between index test and reference 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 two qualified pathologists Classify as 'No' if the biopsy was not independently confirmed by at least two 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 the majority of participants Classify as 'No' if the delay between the index test(s) and reference standard is considered unacceptable for the majority of 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	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	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)

		as 'Unclear' if study did not provide any information on whether interpreters of index tests were blinded to biopsy/histopathology	the study did not provide any information on whether the pathologists were blinded to the index test results	
		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' if 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 prespecified? Classify as 'Yes' if the threshold was prespecified. Classify as 'No' if the threshold was not prespecified. Classify as 'Unclear' if it is unclear whether the threshold was prespecified.		
		Were any conflicts of interest stated? Classify as 'Yes' if the study declared no conflict of interest. Classify as 'No' if the		

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

		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?	

Table 3. Pairwise comparison of diagnostic accuracy

	Vital staining	Cytology	Light-based
Vital staining	Sensitivity 0.84 (0.74 to 0.90) Specificity 0.70 (0.59 to 0.79)		
Cytology	Likelihood-ratio test Sensitivity P = 0.23 Specificity P = 0.003	Sensitivity 0.91 (0.81 to 0.96) Specificity 0.91 (0.81 to 0.95)	
Light-based	Likelihood-ratio test Sensitivity and/or specificity P = 0.49	Likelihood-ratio test Sensitivity P = 0.99 Specificity P = 0.02	Sensitivity 0.91 (0.77 to 0.97) Specificity 0.58 (0.22 to 0.87)

APPENDICES

Appendix I. The Cochrane Oral Health Group Trials Register Search Strategy

An updated search of the Oral Health Group 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 "toluidine b*" or "toluidine b*" or tblue or t-blue or "toluidine dye*" or "toluidine rins*" or "toluidine stain*" or "toluidine 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 orascope 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 "toluidine b*" or "toluidine b*" or tblue or t-blue or "toluidine dye*" or "toluidine rins*" or "toluidine stain*" or "toluidine 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 orascope 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*)))

Appendix 2. Cochrane DTA 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*)

Appendix 3. MEDLINE 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/du
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/du
29. Fluorescence/
30. Spectrometry, fluorescence/
31. exp Luminescent Agents/du
32. Light/du
33. Tomography, Optical Coherence/
34. (visual\$ adj5 (“light emitting diode” or “blue spectrum” or LED or luminous\$)).tw,ot.
35. (visual?ation adj3 adjunct\$).tw,ot.
36. (vizilite or microlux\$ or orascope 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/du
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/
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 4. EMBASE via 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. (visualization adj3 adjunct\$).tw,ot.
34. (vizilite or microlux\$ or orascopic 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. Saliva/
43. Tumor Marker/
44. ((“tumor marker\$” or “neoplas marker\$”) adj3 (blood or saliva)).tw,ot.
45. ((analy\$ or screen\$ or test\$ or examin\$) adj3 (blood or saliva)).tw,ot.
46. Mass screening/
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 5. MEDION Search Strategy

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

CONTRIBUTIONS OF AUTHORS

All review authors collaborated in the conception of the review purpose and design. Drafting the protocol: T Walsh and J Liu. Developing the search strategy: T Walsh and J Liu. Selecting studies for inclusion: R Macey, T Walsh, P Brocklehurst, JLY Liu. Extracting data: R Macey, T Walsh, P Brocklehurst. Carrying out analysis: T Walsh, R Macey. Interpreting the analysis: T Walsh, R Macey. Drafting the final review: R Macey, T Walsh, P Brocklehurst, AR Kerr. Clinical input: P Brocklehurst, AR Kerr, GR Ogden, MW Lingen, S Warnakulasuriya, C Scully. The final review was read and approved by all authors.

DECLARATIONS OF INTEREST

Richard Macey: none known

Tanya Walsh: none known

Paul Brocklehurst: none known

Alexander R Kerr: none known

Joseph LY Liu: 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.

Crispian Scully: none known

SOURCES OF SUPPORT

Internal sources

- MAHSC, UK.

The Cochrane Oral Health Group is supported by the Manchester Academic Health Sciences Centre (MAHSC) and the NIHR Manchester Biomedical Research Centre.

- NIHR-MBRC, UK.

Health Sciences Research, School of Dentistry is supported by the Manchester Academic Health Sciences Centre (MAHSC) and the NIHR Manchester Biomedical Research Centre.

- SDCEP, UK.

This review has been supported by the Scottish Dental Clinical Effectiveness Programme.

External sources

- National Institute for Health Research (NIHR), UK.

The NIHR is the largest single funder of the Cochrane Oral Health Group.

Disclaimer:

The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR, NHS or the Department of Health.

- Cochrane Oral Health Group Global Alliance, Other.

The production of all our reviews is assisted by funding from our Global Alliance partners (<http://ohg.cochrane.org/>): British Association for the Study of Community Dentistry, UK; British Association of Oral Surgeons, UK; British Orthodontic Society, UK; British Society of Paediatric Dentistry, UK; British Society of Periodontology, UK; Canadian Dental Hygienists Association, Canada; Mayo Clinic, USA; National Center for Dental Hygiene Research & Practice, USA; New York University College of Dentistry, USA; and Royal College of Surgeons of Edinburgh, UK

- School of Dentistry, The University of Manchester, UK.

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, and identified these studies in any analyses.

INDEX TERMS

Medical Subject Headings (MeSH)

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

MeSH check words

Humans