Journal of Oral Pathology & Medicine

WILEY

Journal of Oral Pathology & Medicine

REVIEW OPEN ACCESS

Accuracy of Cytological Methods in Early Detection of Oral Squamous Cell Carcinoma and Potentially Malignant Disorders: A Systematic Review and Meta-Analysis

Hoda Tayebi-Hillali[1] | Alejandro I. Lorenzo-Pouso[1,2] | Xabier Marichalar-Mendía[3] | Pilar Gándara-Vila[1,2] | Dolores Reboiras-López[1,2] | Andrés Blanco-Carrión[1,2] | Martina Coppini[4] | Vito Carlo Alberto Caponio[5] | Mario Pérez-Sayáns[1,2,6]

[1]Oral Medicine, Oral Surgery and Implantology Unit (MedOralRes), Faculty of Medicine and Dentistry Universidade de Santiago de Compostela, Santiago de Compostela, Spain | [2]Instituto de Investigación Sanitaria de Santiago (IDIS), ORALRES Group, Santiago de Compostela, Spain | [3]Department of Nursing I, Faculty of Medicine and Nursing, University of Basque Country, Bilbao, Spain | [4]Department of Precision Medicine in Medical, Surgical and Critical Care, University of Palermo, Palermo, Italy | [5]Department of Life Sciences, Health and Health Professions, Link Campus University, Rome, Italy | [6]Instituto de los Materiales de Santiago de Compostela (iMATUS). Avenida Do Mestre Mateo, A Coruña, Spain

Correspondence: Alejandro I. Lorenzo-Pouso (alejandroismael.lorenzo@rai.usc.es)

Received: 18 September 2024 | **Revised:** 16 April 2025 | **Accepted:** 10 June 2025

Funding: The authors received no specific funding for this work.

Keywords: cytology | diagnostic accuracy | leukoplakia | meta-analysis | mouth neoplasm | oral lichen planus | oral potentially malignant disorders

ABSTRACT

Introduction: Oral squamous cell carcinoma (OSCC) carries significant global mortality rates. Brush cytology presents a potential adjunctive tool for early detection and monitoring of OSCC and oral potentially malignant disorders (OPMDs). This study aims to systematically evaluate the diagnostic accuracy of cytology for detecting OSCC and OPMDs compared to histopathology as the reference standard. We conducted a systematic review and meta-analysis following PRISMA-DTA guidelines.

Material and Methods: We searched PubMed, Embase, Scopus, Cochrane Library, and Web of Science from inception to January 2023 (updated in March 2025). Eligible studies included cohort studies evaluating cytology versus histopathological diagnosis. Two reviewers independently screened studies, extracted data, and assessed risk of bias using QUADAS-2. We used the Hierarchical Summary Receiver Operating Characteristic model for meta-analysis.

Results: Of 2603 identified studies, 53 met inclusion criteria, comprising 13,249 patients. Cytology demonstrated a pooled sensitivity of 0.914 (95% CI: 0.878–0.941) and specificity of 0.960 (95% CI: 0.937–0.975). The diagnostic odds ratio was 137.502 (95% CI: 79.733–237.127), with a positive likelihood ratio of 11.970 (95% CI: 9.005–15.912) and negative likelihood ratio of 0.096 (95% CI: 0.059–0.158). Subgroup analysis showed improved performance when exfoliative cytology was combined with DNA analysis or when using a metal spatula. Both conventional and liquid-based cytology were effective, with the latter showing modest advantages. Heterogeneity was substantial across studies ($I^2 = 86.26\%$).

Conclusion: Cytology demonstrates good diagnostic accuracy for detecting OSCC and OPMDs and may serve as a valuable adjunctive screening tool. However, it does not replace histopathological examination as the diagnostic gold standard. Further research should focus on standardizing collection techniques and interpretation criteria. Registration: PROSPERO CRD42023438610.

1 | Introduction

Oral squamous cell carcinoma (OSCC) represents the most common malignant neoplasm in the oral cavity, accounting for over 90% of oral malignancies. It is particularly prevalent in South Central Asia countries, where it ranks among the most frequent types of cancer [1, 2]. OSCC significantly impacts South Asian countries, including Sri Lanka, Pakistan, India, and Bangladesh, largely due to socioeconomic factors and limited awareness of accountable risks like tobacco and alcohol [3]. Despite treatment advances, 5-year survival rates remain below 50%, primarily because diagnoses occur at advanced stages [4]. OSCC may develop from oral potentially malignant disorders (OPMDs) with varying transformation rates, making early detection crucial for preventing cancer progression and treatment complications [5, 6].

Despite the oral cavity being an accessible location to explore for screening purposes, These conditions are often underdiagnosed, leading to delays in the management of these patients. Nowadays, clinical examination followed by a biopsy represents the gold standard. However, there are challenges, as patients may require multiple biopsies to be taken over time during their follow-up, above all for those who are exposed to risk factors [4]. Interest in non-invasive diagnostic methods has grown recently. These tests are increasingly appealing due to their ease of use and painless nature, enabling continuous lesion monitoring.

Thus, while OSCC and OPMDs diagnosis traditionally relies on invasive biopsies requiring specialized skills, often causing referral delays [7], cytology offers a minimally invasive alternative for analyzing oral mucosa cells. This technique involves cell collection through scraping specialized instruments like Cytobrush, Oral CDx, or dermatological curettes [8, 9]. In addition, advanced liquid-based cytology (LBC) is a modern diagnostic technique that involves collecting cell samples in a liquid medium, which enhances cell preservation and ensures more uniform sample preparation. This method provides high-quality specimens that are ideal for various analyses, including immunohistochemical testing, enabling more accurate and reliable results. LBC offers several advantages over traditional smear methods, such as reducing sample contamination and increasing sensitivity [10].

Although two previous systematic reviews have addressed exfoliative cytology in oral cancer diagnosis, significant limitations persist. One review did not conduct a pooled analysis of diagnostic accuracy, while the other focused exclusively on brush cytology techniques without adhering to rigorous methodological guidelines, and did not follow the PRISMA-DTA guidelines [10, 11]. Our systematic review and meta-analysis employs more comprehensive inclusion criteria and exhaustive search strategies, resulting in a substantially larger sample size and implementing more robust statistical approaches to provide definitive evidence on this diagnostic method. In this vein, this systematic review and meta-analysis aims to comprehensively assess the accuracy of cytological methods for early diagnosis of OSCC and OPMDs.

2 | Materials and Methods

The protocol for this meta-analysis and systematic review was properly registered in The International Database of Prospectively Registered Systematic Reviews (PROSPERO) under the protocol number CRD42023438610. The study strictly adhered to the PRISMA-DTA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Diagnostic Test Accuracy) guidelines [12], which provided a comprehensive framework specifically designed for conducting and reporting diagnostic test accuracy reviews, ensuring methodological rigor, transparency, and reproducibility in our evaluation of cytological techniques.

2.1 | Data Sources

Studies relevant to this topic were identified through a comprehensive search conducted across multiple databases, including PubMed, Embase, Scopus, Cochrane Library, and Web of Science. All studies up to January 2024 were included, with a subsequent update in March 2025.

The following search syntax was applied across all databases using synonyms and free terms to maximize sensitivity: ("oral" OR "mouth" OR "buccal" OR "oral cavity" OR "oral mucosa" OR "lip" OR "lips" OR "tongue" OR "gingiva" OR "palate" OR "cheek" OR "intra-oral" OR "intraoral" OR "gum" OR "gums" OR "labial") AND ("tumor" OR "cancer" OR "carcinoma" OR "carcinogen" OR "neoplasm" OR "malignant" OR "metastasis" OR "dysplasia" OR "lesion" OR "ulcer" OR "precancer" OR "precancer" OR "premalignant" OR "precursor" OR "lichen planus" OR "leukoplakia" OR "submucous fibrosis" OR "actinic keratosis" OR "candidiasis" OR "erythroplakia" OR "erythroplasia" OR "erythroleukoplakia" OR "hyperplasia" OR "hyperkeratosis") AND ("cytodiagnosis" OR "cytophotometry" OR "brush biopsy" OR "oral cdx" OR "oralcdx" OR "modified liquid based cytology" OR "exfoliative cytology").

The detailed search syntax for all additional databases can be found in the Supporting Information (Appendix S1).

2.2 | Eligibility Criteria and Study Selection Process

Diagnostic cohort studies evaluating the accuracy of oral cytology techniques in detecting OSCC or OPMDs were included. Eligible studies had to:

1. Assess individuals with OSCC (AJCC codes: C00-C06) or OPMDs, following WHO diagnostic criteria [5, 13]. If studies did not explicitly use these criteria, they were included if they applied reliable diagnostic frameworks based on previous classifications or provided exhaustive descriptions aligning with them, as previously described [14].
2. Use histopathological diagnosis from biopsies as the reference standard for suspicious lesions.
3. Report true positive (TP), false positive (FP), false negative (FN), and true negative (TN) values or provide sufficient data for calculation.
4. Have a minimum sample size of 20.
5. Be published in any language or year.

The following references were excluded:

1. Studies on recurrent neoplasms.
2. Retracted articles, letters, editorials, opinion pieces, comments.
3. Cross-sectional, in vitro, and non-human studies.

The deduplicated list of records from the merged databases was independently screened by two authors (H.T.H. & A.I.L.P.) using EndNote 21 (Clarivate Analytics, Philadelphia, PA, US). Initial screening was based on titles and abstracts, yielding a selection of articles for full-text eligibility assessment. Inter-reviewer agreement was evaluated at this stage using the kappa (κ) coefficient, and any discrepancies were resolved through discussion with a third author (M.P.S.). The final selection of studies was compiled in accordance with the predefined inclusion and exclusion criteria by H.T.H. & A.I.L.P.

2.3 | Focused PICO Question

(P) Participants: Patients with OSCC (AJCC codes: C00-C06) or OPMDs, diagnosed based on WHO criteria or, if not explicitly stated, using reliable classifications or exhaustive descriptions analogous to them [5, 13].

(E) Intervention: Use of cytology (including brush and liquid-based cytology) for OSCC and OPMD detection.

(C) Comparator: Histopathological diagnosis from biopsies as the reference standard of suspicious lesions.

(O) Outcome: Diagnostic accuracy measures, including sensitivity, specificity, diagnostic odds ratio (DOR), and likelihood ratios (PLR and NLR).

2.4 | Study Variables

Two investigators (H.T.H. & A.I.L.P.) independently extracted data using a standardized form for each eligible study. Variables extracted from each primary-level study included the lead author, year of publication, country, setting or organization in which the study was conducted, study type (prospective or retrospective), and the sex and age of participating patients. Information on lesion locations and risk factors associated with participants was also collected. Regarding samples, data on the number per study, lesions related to OSCC or OPMDs, and the technique and instrument used for sample collection were gathered. In cases where the information was incomplete, attempts were made to contact the authors to request the missing information. For the diagnostic performance analysis, values for TP, FP, TN, and FN were collected.

2.5 | Quality Assessment of Individual Studies

To diminish the risk of bias, two distinct reviewers (M.C. & M.P.S.) analyzed the selected articles. The quality of included studies was assessed using the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) checklist [15]. The QUADAS-2 checklist was modified following Whiting et al.'s (2011) recommendations. Since the methodologies evaluated in the review were quantified using an "objective" approach, the question in domain 2 regarding blinding of the test interpreter to the reference standard results was omitted [16]. Study quality was categorized as high (6–7 points), moderate (4–5 points), or low (0–3 points) for meta-analysis.

2.6 | Quantitative Synthesis

For the meta-analysis, the Hierarchical Summary Receiver Operating Characteristic (HSROC) model was employed. The HSROC model directly estimates parameters, facilitating the construction of an HSROC curve including summarized points of sensitivity (S) and specificity (E), along with their prediction and confidence region. Globally, we calculated the Diagnostic Odds Ratio (DOR), S, E, Positive Likelihood Ratio (PLR), and Negative Likelihood Ratio (NLR) using the DerSimonian-Laird random-effects estimator or Mantel–Haenszel for fixed-effects. The *ad hoc* interpretation of Area Under the Curve (AUC) values was as follows: >0.97 "excellent", 0.93–0.96 "very good", 0.75–0.92 "good", and 0.5–0.75 "fair" [17].

To assess heterogeneity, the I^2 test was conducted using the formula $I^2 = 100\times\{[Q-(k-1)]/Q\}$, where Q represents Cochran's Q, a $\chi 2$ test with (k–1) degrees of freedom, and k is the number of included studies. Based on the I^2 statistics, heterogeneity was categorized as low ($I^2 < 50\%$) or high ($I^2 > 50\%$). Egger's test was used to analyze publication bias, with $p_{Egger} < 0.1$ considered significant [18, 19]. Funnel plots following Deeks' method were visually inspected for bias. A bivariate boxplot was also generated to assess outliers and describe diagnostic value. The Fagan diagram evaluated the clinical utility of a diagnostic test by comparing the pre-test probability with the post-test probability, using likelihood ratios as a connection between them. A p-value < 0.05 was considered statistically significant. Statistics were performed using the open-source statistical programming language R (v4.1.2; R Core Team, 2021) and Stata 16.1 (Stata Corp, College Station, TX, USA).

3 | Results

3.1 | Qualitative Results

A total of 2603 articles were identified, out of which 53 were included as primary-level studies for our analysis. Several references were excluded due to non-compliance with eligibility criteria, including narrative reviews, cross-sectional studies, or inclusion of samples from entirely healthy subjects. Others were excluded due to unclear and irrelevant information (Figure S1). The included studies were published between 1960 and 2024. The global agreement between the two reviewers was 90.68% and the Kappa score was 0.9 for the included studies (full-text articles), indicating substantial agreement.

Most studies were conducted in Europe, particularly Germany, and India in Asia. No studies from Africa were found. Among the 53 studies included, 77.36% were prospective ($n = 41$) and 22.64% retrospective ($n = 12$). On the other hand, our final analysis included

67 units of analysis. This discrepancy arises because several studies provided multiple datasets, either by comparing different collection instruments, evaluating distinct cytological techniques, or analyzing separate lesion types within the same publication. Each of these distinct comparisons was treated as an individual analytical unit to enable more granular analysis and maximize the information extracted from the primary-level studies.

The male-to-female ratio is similar, with a moderate male predominance (M:F = 1.78), and the average age of subjects is 45–65 years. Tobacco and alcohol were major risk factors, with chewing tobacco prevalent in India and Sri Lanka. Lesions commonly appear on the tongue's lateral aspect and oral mucosa, less frequently on the labial mucosa, mouth floor, and palate. Among the suspicious lesions, a substantial cohort of 3181 cases have been classified as OSCC. Additionally, a significant group of 3905 cases have been identified as OPMDs. A detailed characterization of this latter category reveals the presence of 2347 cases of leukoplakia, 1222 cases of lichenoid conditions, 119 cases of oral submucous fibrosis, 36 cases of erythroleukoplakia, 66 cases of proliferative verrucous leukoplakia, and 11 cases of palatal lesions associated with reverse smoking.

For sample collection, conventional cytology was predominantly used in most studies. Others, however, combined conventional cytology with DNA image cytometry, or with galectin-1 analysis, AgNOR, toluidine blue, immunohistochemistry of laminin-5 γ2 chain, or microRNA. On the other hand, liquid-based cytology was employed in 14 studies. In terms of instruments used, the Cytobrush and its Cytobrush Plus GT version stood out. The Orecellex brush was also utilized in several studies. Only three studies used tongue depressors for sample collection, and another twelve used nylon toothbrushes. Metal spatulas were used in only two studies. The data extracted from each article is depicted in Table 1. All primary-level-studies are reflected in Appendix S2.

3.1.1 | Quality Assessment of the Included Studies

In patient selection, 40 studies were evaluated with low risk of bias, although ten showed high risk by including healthy controls and inappropriate exclusions. Regarding diagnostic testing, 35 studies were low risk, but 12 had inadequate samples, increasing bias risk. All studies used biopsy as the gold standard, with 88.68% classified as low risk, except for two with inadequate biopsies. In the flow and timing domain, 30% showed high risk due to lack of corresponding biopsies. Concerning the delay time between diagnostic test and gold standard, a three-week interval was mentioned in Seijas-Naya et al.'s study, and one month in Scheifele et al.'s (Figures S2 and S3). As a result, 3 studies were rated as "low" quality, 20 as "moderate," and 30 as "high."

3.2 | Quantitative Results

3.2.1 | Diagnostic Accuracy

The total sample size included in this meta-analysis was 4132 adult patients with OSCC and OPMDs. The overall S and E were 0.914 (95% confidence interval, CI: 0.878–0.941) and 0.960 (95% CI: 0.937–0.975), respectively. The total diagnostic odds ratio (DOR) was 137.502 (95% CI: 79.733–237.127). The PLR was 0.119 (95% CI: 0.090–0.159) and the NLR was 0.096 (95% CI: 0.059–0.158). Additionally, the SROC curve and the AUC were plotted. The overall AUC was 0.980 (95% CI: 0.960–0.990), indicating that cytology demonstrates "excellent" accuracy in diagnosing these lesions (Figure 1).

3.2.2 | Heterogeneity and Subgroup Analysis

The studies included in the meta-analysis showed substantial heterogeneity ($I^2 = 86.26\%$), as depicted in Figure 2. To investigate the sources of heterogeneity, a subgroup analysis was conducted, considering several potential predictors: continent of the study, study type, technique used, sample collection instrument, diagnosed lesion, sample size, and study quality assessed by QUADAS-2 (Table 2). Despite these efforts, all subgroups displayed high heterogeneity. Briefly, the main results based on subgroups are displayed below:

i. Diagnostic performance by technique used

Conventional cytology followed with a sensitivity of 0.874 (95% CI: 0.826–0.910), while liquid-based cytology showed a sensitivity of 0.859 (95% CI: 0.742–0.928). Among the various diagnostic techniques, cytology combined with immunohistochemistry or microRNA analysis demonstrated the highest sensitivity (0.919; 95% CI: 0.848–0.958).

ii. Sample collection instrument

The instrument used for sample collection also significantly influenced diagnostic performance. The metal spatula, used in two studies, achieved the highest sensitivity (0.928; 95% CI: 0.889–0.955) and specificity (0.988; 95% CI: 0.967–0.996). The tongue depressor demonstrated high sensitivity (0.924; 95% CI: 0.789–0.975), but the lowest specificity (0.832; 95% CI: 0.338–0.980).

iii. Diagnostic performance by lesion type

Exfoliative cytology showed higher sensitivity in diagnosing OSCC (0.911; 95% CI: 0.855–0.947). However, its sensitivity decreased significantly when diagnosing OPMDs (0.724; 95% CI: 0.405–0.910), resulting in more false negatives. Conversely, the specificity was slightly higher in diagnosing OPMDs (0.966; 95% CI: 0.904–0.988).

iv. Performance by study continent

Analysis by continent showed that studies conducted in America exhibited the highest sensitivity (0.966; 95% CI: 0.930–0.984) and a high specificity (0.927; 95% CI: 0.892–0.952). However, the highest specificity was observed in studies conducted in Oceania (0.965; 95% CI: 0.907–0.987).

v. Sample size and study quality

Studies with more than 100 samples showed higher sensitivity (0.909; 95% CI: 0.862–0.941) and specificity (0.931; 95% CI: 0.899–0.953). Sensitivity also increased as the quality of the studies, as assessed by QUADAS-2, improved. No significant differences were found between prospective and retrospective studies.

TABLE 1 | Data on location and type of study; sex and age of the patients; number of samples; type of lesion; risk factor consumed by the patients; location of lesions; type of cytology; instrument and the sensitivities and specificities of each study.

Author and year	Country	Study type	Sex and age	Samples	Lesions	Risk factors
Cahn et al., 1959	USA	Prospective	Not specified	239 samples from 239 patients	OSCC (131)	Not specified
Shklar et al.,1968	USA	Prospective	927 women, 1125 men (> 45 years)	2052 samples from 2052 patients	OSCC (83); leukoplakia (250); OLP (36); lymphosarcoma (3); melanoma (1)	Not specified
Dabelesteen et al., 1979	Denmark	Prospective	Not specified	299 samples from 269 patients	Leukoplakia (299)	Not specified
Sciubba, 1999	Baltimore	Prospective	502 women, 443 men (18–83 years average)	945 samples from 945 patients	OSCC+ OPMDs (131)	Tobacco and alcohol
Remmerbach et al., 2001	Germany	Prospective	43% women, 57% men (60 years average)	1254 samples from 181 patients	OSCC (56); OLP (49); leukoplakia (83)	Not specified
Remmerbach et al., 2003	Germany	Prospective	25.3% women, 74.7% men (55.9 years average)	337 samples from 75 patients	OSCC (53); leukoplakia (22)	Not specified
Scheifele et al., 2004	Germany	Prospective	41.3% women (64.3 years), 58.85 men (53.2 years average)	103 samples from 80 patients	OSCC (13); leukoplakia (49); OLP (18)	Not specified
Poate et al.,2004	United Kingdom	Retrospective	50 men, 62 women (55 years average)	112 samples from 112 patients	OSCC (7)	Tobacco, pipe and alcohol
Maraki et al., 2004	Germany	Prospective	54% women, 46% men (61 years average)	98 samples from 98 patients	OSCC (15); leukoplakia (21); erythroplakia (3); OLP (37)	Not specified
Hayama et al., 2004	Brazil	Retrospective	29 women (49.3 años), 15 men (50.7 years average)	44 samples from 44 patients	OSCC (11); OLP (8); leukoplakia (2)	Tobacco and alcohol
Brunotto et al.,2005	Argentina	Prospective	(51–68 years average)	46 samples from 46 patients	OSCC (3); OLP (4); leukoplakia (2)	Not specified
Maraki et al., 2006	Germany	Prospective	72% women, 28% men (58 years average)	56 samples from 56 patients	OSCC (2); OLP (4)	Not specified
Driemel et al.,2007	Germany	Retrospective	Not indicated	93 samples from 93 patients	OSCC (27)	Not specified
Driemel et al.,2007	Germany	Prospective	Not specified	159 samples from 159 patients	OSCC (52)	Not specified

(Continues)

TABLE 1 | (Continued)

Author and year	Country	Study type	Sex and age	Samples	Lesions	Risk factors
Navone et al., 2007	Italy	Prospective	Not specified	384 samples from 384 patients	OSCC (64)	Not specified
				89 samples from 89 patients	OSCC (32)	Not specified
Mehrotra et al., 2007	India	Prospective	25% women 75% men (52 years average)	96 samples from 96 patients	OSCC (32); oral submucous fibrosis (45); leukoplakia (10)	Smoked and chewed tobacco, chewed paan, and alcohol
Mehrotra et al., 2008	India	Prospective	76% men 24% women (50 years average)	79 samples from 79 patients	OSCC (28); leukoplakia, erythroplakia and oral submucous fibrosis (45)	Tobacco and alcohol
Majert et al., 2009	Germany	Prospective	67% men, 33% women (68.9 años media)	69 samples from 69 patients	OSCC (15)	Not specified
Remmerbach et al., 2009	Germany	Prospective	(61 years average)	47 samples from 47 patients	OSCC (20); leukoplakia (7); OLP (20)	Not specified
Rajput et al., 2010	India	Prospective	Not specified	88 samples from 44 patients	OSCC (34)	Not specified
Delavarian et al., 2010	Iran	Prospective	12 women, 13 men (54 years average)	26 samples from 25 patients	OSCC (12); leukoplakia (5); OLP (7);	Not specified
Güneri et al., 2011	Turkey	Prospective	13 men, 22 women (56.2 years average)	43 samples from 35 patients	OSCC (12); OLP (7); adenocarcinoma (1); lichenoid lesion (2)	Trauma, chronic inflammation, and tobacco
Mehrotra et al.,2011	India	Prospective	55 men, 30 women (45.5 years average)	85 samples from 85 patients	OSCC + OPMDs (27)	Tobacco and alcohol
Babshet et al., 2011	India	Prospective	56 men, 4 women (45–56 years average)	60 samples from 60 patients	OSCC (30); OPMDs (30)	Tobacco, gutka, betel nut, and alcohol
Koch et al., 2011	Germany	Prospective	(62.8 years average)	182 samples from 135 patients	OSCC (104); OLP (8)	Not specified
Seijas-Naya et al., 2012	Spain	Prospective	12 women, 12 men (40–82 years average)	24 samples from 24 patients	Leukoplakia (19); erythroleukoplakia (5)	Tobacco and alcohol
Pérez-Sayáns et al., 2012	Spain	Prospective	29 men, 19 women (44–67.65 years average)	48 samples from 48 patients	OSCC (28); leukoplakia (12); OLP (8)	Tobacco
Rahman et al., 2012	India	Prospective	68 men, 18 women (43 years average)	107 samples from 86 patients	OSCC (28)	Tobacco (smoked and chewed) and alcohol

(Continues)

TABLE 1 | (Continued)

Author and year	Country	Study type	Sex and age	Samples	Lesions	Risk factors
Ng et al., 2012	Canada	Retrospective	82 men, 89 women (58 years average)	171 samples from 171 patients	OSCC (14); OPMDs (11)	Tobacco
Kämmerer et al., 2013	Germany	Prospective	45 men, 25 women (62 years average)	88 samples from 70 patients	OSCC (25); OLP (15)	Not specified
Fontes et al., 2013	Brazil	Retrospective	114 men, 58 women (20–93 years average)	172 samples from 172 patients	OSCC (156); OLP (1)	Not specified
Gupta et al., 2014	India	Prospective	175 men, 27 women (30–60 years average)	117 samples from 202 patients	OSCC (79); oral submucous fibrosis (72); leukoplakia (56); verrucous leukoplakia (8); lesions from reverse smoking (8); erosive OLP (1)	Tobacco and gutka
Ma et al., 2014	China	Prospective	37 women, 15 men (58 years average)	52 samples from 52 patients	OLP (21); leukoplakia (6)	Not specified
Trakroo et al., 2015	India	Prospective	43 women, 7 men (20–70 years average)	50 samples from 50 patients	OSCC (17), OPMDs (33)	Tobacco and alcohol
Kaur et al., 2016	India	Prospective	78 men, 22 women (51–60 years average)	100 samples from 100 patients	OSCC (49)	Tobacco (smoked and chewed), betel, and alcohol
Nanayakkara et al., 2016	Sri Lanka	Prospective	149 men, 43 women (21–95 years average)	192 samples from 192 patients	OSCC (69); leukoplakia (112)	Smoked tobacco, betel and alcohol
Noda et al., 2016	Japan	Prospective	15 women, 22 men (68 years average)	37 samples from 37 patients	OSCC (16)	Not specified
He et al., 2016	USA	Retrospective	Not indicated	39 samples from 39 patients	OSCC (19)	Not specified
Remmerbach et al., 2017	Germany	Prospective	72 men, 41 women (66.5 years average)	113 samples from 113 patients	OSCC (81)	Not specified
Jajodia et al., 2017	India	Prospective	15 women, 33 men (53.45 years average)	48 samples from 48 patients	OSCC (31); leukoplakia (13); erythroleukoplakia (4)	Betel, areca nut, and alcohol
Skandarajah et al., 2017	India	Retrospective	8 women, 24 men (50 years average)	32 samples from 32 patients	OSCC (19); lymphoma (1); erythroplakia (1); leukoplakia (5); proliferative verrucous leukoplakia (1); verrucous leukoplakia (2); tobacco-related lesions (3); OLP (1)	Alcohol, smoked and chewed tobacco

(Continues)

TABLE 1 | (Continued)

Author and year	Country	Study type	Sex and age	Samples	Lesions	Risk factors
Goodson et al., 2017	Australia	Retrospective	170 men, 140 women (18–91 years average)	310 samples from 310 patients	Leukoplakia (268); erythroleukoplakia (27); erythroplakia (15)	Not specified
Kujan et al., 2018	Australia	Prospective	46 women, 40 men (62 years average)	114 samples from 86 patients	OSCC (11); OLP (6)	Not specified
Liu et al., 2019	China	Retrospective	98 men, 105 women (53 years average)	203 samples from 203 patients	OSCC (42); leukoplakia (68); verrucous leukoplakia (1); OLP (83)	Tobacco and alcohol
Remmerbach et al., 2019	Germany	Retrospective	552 women, 440 men (61.6 years average)	1352 samples from 992 patients	OSCC (105); leukoplakia (297); PVL (20); OLP (260); erosive OLP (138); erythroplakia (29)	Not specified
Adhya et al., 2019	India	Prospective	(58 years average)	280 samples from 280 patients	OSCC (221)	Not specified
Remmerbach et al., 2021	Germany	Retrospective	52% women, 48% men (63 years average)	2018 samples from 2018 patients	OSCC (181); OLP (524); leukoplakia (421); PVL (33); erythroplakia (34)	Not specified
Kujan et al., 2021	Australia	Prospective	39 men, 33 women (64.75 years average)	72 samples from 72 patients	Leukoplakia (12%); OLP (5%); oral submucosal fibrosis (2%)	Not specified
Neumann et al., 2022	Germany	Retrospective	315 men (59.0 years average); 355 women (64.1 years average)	814 samples from 670 patients	OSCC (74); leukoplakia (232); PVL (1); erythroplakia (22); OLP (242)	Not specified
Bechstedt et al., 2022	Germany	Prospective	380 women (57.7 years), 222 men (63.65 years)	602 samples from 467 patients	OSCC (308)	Not specified
Kujan et al., 2022	Australia	Prospective	149 women, 135 men (63.24 years average)	284 samples from 284 patients	OSCC (19) OLP (57)	Not specified Not specified
Kokubun et al., 2023	Japan	Prospective	299 men (58.6 years average), 354 women (60.4 years average)	653 samples from 653 patients	OSCC (47); erythroplakia (7)	Not specified
Liu et al., 2024	Canada	Prospective	117 men, 64 women (61 ± 14 years)	190 samples from 181 patients	OSCC (95); normal (95)	Smoked tobacco

Author and year	Location	Technique used	Instrument	Sensitivity	Specificity
Cahn et al., 1959	Not specified	Conventional cytology	Tongue depressor	89.70%	59.10%
Shklar et al.,1968	Not specified	Conventional cytology	Tongue depressor	97.80%	93.30%

(Continues)

TABLE 1 | (Continued)

Author and year	Location	Technique used	Instrument	Sensitivity	Specificity
Dabelesteen et al., 1979	Not specified	Conventional cytology	Not specified	37.50%	98.50%
Sciubba, 1999	Lateral tongue (143); ventral tongue (58); floor of mouth (54); buccal mucosa (243); hard palate (54); soft palate (21); retromolar trigone (26); oropharynx (3); attached gingiva (153); labial and alveolar mucosa (116)	Conventional cytology	OralCDx	100%	92.80%
Remmerbach et al., 2001	Not specified	Conventional cytology	Cytobrush plus GT	94.60%	99.50%
		Conventional cytology combined with DNA image cytometry	Cytobrush plus GT	98.20%	100%
Remmerbach et al., 2003	Floor of mouth (6); tongue (11); tongue and floor of mouth (12); lip (0); uvula and palate (10); buccal mucosa (7); alveolar ridge (7)	Conventional cytology	Cytobrush plus GT	92.50%	100%
Scheifele et al., 2004	Not specified	Conventional cytology	OralCDx	92.30%	94.30%
Poate et al.,2004	Floor of mouth (9); lateral tongue (23); ventral tongue (7); lip (7); palate (17); buccal mucosa (25); alveolar mucosa (3); retromolar trigone (3)	Conventional cytology	OralCDx	71.40%	32%
Maraki et al., 2004	Buccal mucosa (10); tongue (10); lower lip (4); alveolar mucosa (2); retromolar trigone (1)	Conventional cytology combined with DNA image cytometry	Cytobrush plus GT	100%	97.40%
Hayama et al., 2004	Not specified	Conventional cytology	Cytobrush	100%	100%
		Liquid-based cytology	Cytobrush	100%	100%

(Continues)

TABLE 1 | (Continued)

Author and year	Location	Technique used	Instrument	Sensitivity	Specificity
Brunotto et al., 2005	Not specified	Conventional cytology (OPMD)	Cytobrush	100%	100%
		Conventional cytology (OSCC)	Cytobrush	100%	100%
Maraki et al., 2006	Not specified	Conventional cytology combined with DNA image cytometry	Cytobrush plus GT	100%	92.50%
Driemel et al.,2007	Floor of mouth (13); tongue (5); soft palate (1); buccal mucosa (1); alveolar process and retromolar trigone (7)	Conventional cytology + laminin γ2 chain	Cytobrush Plus GT	100%	97%
Driemel et al.,2007		Conventional cytology	Cytobrush Plus GT	78%	96%
	Not specified	Conventional cytology + high molecular weight tenascin C	Cytobrush Plus GT	95%	99%
Navone et al., 2007	Not specified	Liquid-based cytology	Metal spatula	95.10%	99.00%
		Conventional cytology	Cytobrush	85.70%	95.90%
Mehrotra et al., 2007	Not specified	Conventional cytology	Nylon toothbrush	95.60%	94%
Mehrotra et al.,2008	Buccal mucosa (10); tongue (6); lower lip (5); alveolar mucosa (4); floor of mouth (2)	Conventional cytology	Nylon toothbrush	88.23%	93.30%
Majert et al., 2009	Tongue (26); buccal mucosa (18); palate (9); lip (4); floor of mouth (18); alveolar mucosa (28); retromolar trigone (7)	Conventional cytology	OralCDx	52%	29%

(Continues)

TABLE 1 | (Continued)

Author and year	Location	Technique used	Instrument	Sensitivity	Specificity
Remmerbach et al., 2009	Not specified	Conventional cytology	Cytobrush	100%	92.60%
	Not specified	Conventional cytology with DNA image cytometry	Cytobrush	90%	100%
		Conventional cytology with AgNOR analysis	Cytobrush	100%	100%
Rajput et al., 2010	Not specified	Conventional cytology with AgNOR analysis	Cytobrush Plus GT	100%	100%
Delavarian et al., 2010	Not specified	Liquid-based cytology	OralCDx	88.80%	100%
Güneri et al., 2011	Buccal mucosa (56%); tongue (19%); hard palate (14%)	Conventional cytology	Cytobrush Plus GT	92.30%	51.70%
Mehrotra et al.,2011	Buccal mucosa (38); tongue and floor of mouth (18); lip and alveolar mucosa (8); hard palate (7); gingiva (8)	Conventional cytology	OralCDx	96.30%	90.40%
Babshet et al., 2011	Buccal mucosa (40); tongue (3); lip (4); alveolus or gingiva (13)	Conventional cytology	Nylon toothbrush	77%	100%
Koch et al., 2011	Not specified	Conventional cytology	Cytobrush Plus GT	93.9%	94%
Seijas-Naya et al., 2012	Buccal mucosa (4); tongue (14); lip (1); retromolar trigone (2); hard palate (1); gingiva (2)	Conventional cytology	OralCDx	72.7%	92.3%
Pérez-Sayáns et al., 2012	Not specified	Conventional cytology	Cytobrush	69%	100%
Rahman et al., 2012	Not specified	Conventional cytology	Cytobrush	70.37%	77.96%
Ng et al., 2012	Tongue (68); floor of mouth (9); soft palate and retromolar trigone (23); buccal mucosa (57); gingiva and hard palate (14)	Liquid-based cytology	Nylon toothbrush	89%	97%

(Continues)

TABLE 1 | (Continued)

Author and year	Location	Technique used	Instrument	Sensitivity	Specificity
Kämmerer et al., 2013	Not specified	Conventional cytology	Cytobrush Plus GT	89%	100%
		Conventional cytology with DNA image cytometry	Cytobrush Plus GT	70%	100%
	Buccal mucosa (160); tongue (15)	Conventional cytology with DNA image cytometry	Cytobrush Plus GT	77%	100%
Fontes et al., 2013	Not specified	Conventional cytology	Cytobrush	99.3%	100%
Gupta et al., 2014	Buccal mucosa (160); tongue (15)	Conventional cytology	Tongue depressor	82.8%	33.3%
		Conventional cytology	Nylon toothbrush	87.1%	52.9%
Ma et al., 2014	Not specified	Conventional cytology with DNA image cytometry	Cytobrush	86.36%	90%
Trakroo et al., 2015	Tongue (2); tuberosity (1); buccal mucosa (28); palate (1); gingiva (2)	Conventional cytology	Cytobrush	84.37%	88.89%
Kaur et al., 2016	Buccal mucosa (48%); lateral border of tongue (34%); lower lip (4%); alveolus (6%); hard palate (3%); angle of mouth (4%); and floor of mouth (1%)	Conventional cytology	Nylon toothbrush	83.3%	95.8%
		Conventional cytology with DNA image cytometry	Nylon toothbrush	68.7%	100%
Nanayakkara et al., 2016	Tongue (20); buccal mucosa (101); lip (14); alveolus (17); palate (7); others (33)	Conventional cytology	Cytobrush Plus GT	98.8%	100%
		Conventional cytology	Metal spatula	92.2%	100%
Noda et al., 2016	Tongue (11); gingiva (19); floor of mouth (1)	Liquid-based cytology + Gal1 analysis	Cytobrush	75%	66.6%
He et al., 2016	Tongue (19)	Conventional cytology + microRNA analysis	Cytobrush	100%	64%

(Continues)

TABLE 1 | (Continued)

Author and year	Location	Technique used	Instrument	Sensitivity	Specificity
Remmerbach et al., 2017	Not specified	Conventional cytology	Nylon toothbrush	96.3%	90.63%
		Liquid-based cytology	Nylon toothbrush	97.53%	68.75%
Jajodia et al., 2017	2/3 anterior of tongue (5); anterior pillar of tonsil (1); buccal mucosa (25); floor of mouth (2); retromolar trigone (2); alveolus (9); hard palate (3); lip (1)	Conventional cytology	Nylon toothbrush	84%	50%
		Liquid-based cytology	Nylon toothbrush	75%	50%
Skandarajah et al., 2017	Buccal mucosa (21); tongue (7); lip (2); retromolar trigone (1); alveolus (1)	Liquid-based cytology	Cytobrush Plus GT	70%	100%
Goodson et al., 2017	Floor of mouth (62); lateral tongue (60); dorsal tongue (13); ventral tongue (17); palate (38); buccal mucosa (55); alveolar process (24); lip (11); retromolar trigone (7); labial commissure (5)	Conventional cytology	Orcellex brush	60%	99%
Kujan et al., 2018	Lateral and ventral tongue (38); dorsal tongue (1); palate (10); buccal mucosa (21); lip mucosa and commissure (3); gingiva (4); retromolar trigone (4); alveolar mucosa (7); floor of mouth (11); lower lip vermilion (2)	Liquid-based cytology	Orcellex brush	75%	76%
Liu et al., 2019	Tongue (44); buccal mucosa (22); gingiva (4)	Conventional cytology	Cytobrush	79%	81%
Remmerbach et al., 2019	Alveolar ridge (30.3%); buccal mucosa (28.5%); lateral border of tongue (22.4%)	Liquid-based cytology	Orcellex brush	95.6%	84.9%

(Continues)

TABLE 1 | (Continued)

Author and year	Location	Technique used	Instrument	Sensitivity	Specificity
Adhya et al., 2019	Tongue (21%); gingivobuccal sulcus (21%); buccal mucosa (20%); retromolar trigone (15%); lower alveolar mucosa (13%); upper alveolar mucosa (10%)	Conventional cytology	Not specified	98.2%	89.3%
Remmerbach et al., 2021	Not specified	Liquid-based cytology	Orcellex brush y Cytobrush plus GT	92.8%	99.2%
Kujan et al., 2021	Lateral tongue (3); palate (9); floor of mouth (4); lip mucosa and commissure (3); retromolar trigone (4)	Liquid-based cytology	Orcellex brush	91.1%	92.8%
Neumann et al., 2022	Not specified	Liquid-based cytology	Orcellex brush	100%	86.2%
Bechstedt et al., 2022	Tongue (30.3%); floor of mouth (21.54%); buccal mucosa (17.7%)	Conventional cytology	Cytobrush plus GT	100%	69.4%
Kujan et al., 2022	Labial and buccal mucosa (97); lateral tongue (77); floor of mouth (45); alveolar ridge and palate (65)	Liquid-based cytology	Orcellex brush	89.47%	99.25%
		Liquid-based cytology	Orcellex brush	75.38%	96.35%
Kokubun et al., 2023	Tongue and floor of mouth (18); gingiva (8); buccal mucosa (38); palate (7); lip (8); floor of mouth (1)	Liquid-based cytology	Orcellex brush	69%	75%
Liu et al., 2024	Buccal mucosa, palate, gingiva (67); lengua, floor of mouth (123)	Conventional cytology + DNA ploidy	Nylon toothbrush	98.9%	90.5%

Abbreviations: OLP, Oral Lichen Planus; OSCC, Oral Squamous Cell Carcinoma; PVL, Proliferative Verrucous Leukoplakia.

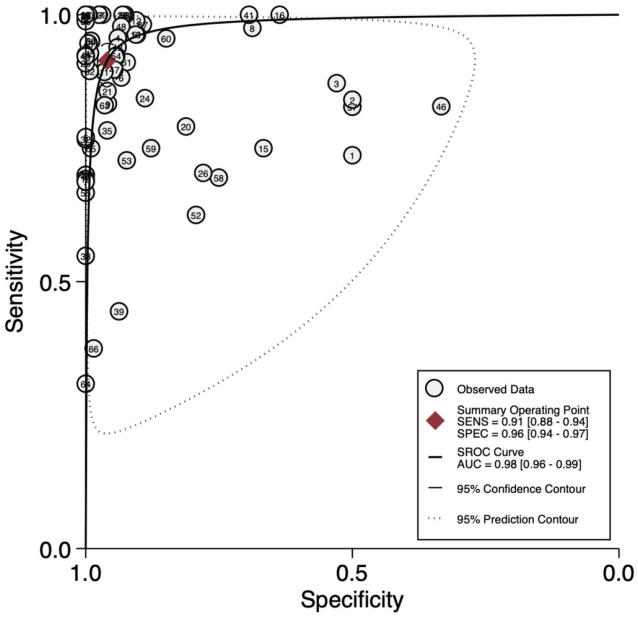

FIGURE 1 | Curve SROC. Summary ROC curves for 67 statistical unit studies reporting on diagnostic accuracy of cytology for oral squamous cell carcinoma and potentially malignant disorders.

3.2.3 | Complementary Analysis

In Figure 3, a symmetry is observed in the test performance measures with similar sensitivity and specificity. The graph also displayed the heterogeneity among the different primary-level studies, with 10 outliers. Therefore, for quantitative synthesis, the use of a bivariate random-effects model was necessary. In terms of "*small study effects*", Egger's test was used to assess publication bias among the primary-level studies, yielding a P_{Egger} value of 0.02, indicating significant publication bias, which was confirmed formally by a funnel plot (Figure 4a).

When addressing the clinical utility of the test, the probability of having OSCC and OPMDs prior to performing exfoliative cytology is 25%, increasing to 88% after a positive test result. Additionally, the probability of not having OSCC or OPMDs after the test is 3%, indicating that exfoliative cytology exhibits high specificity and can effectively rule out the presence of these lesions (Figure 4b).

4 | Discussion

This meta-analysis provides a comprehensive evaluation of cytological methods for the early detection of OSCC and OPMDs. Our findings reveal that exfoliative cytology demonstrates excellent diagnostic accuracy with an AUC of 0.980 (95% CI: 0.960–0.990), with overall sensitivity and specificity of 0.914 (95% CI: 0.878–0.941) and 0.960 (95% CI: 0.937–0.975). These results align with previous meta-analyses such as Dolens et al., who reported similar high sensitivity (0.942; 95% CI: 0.926–0.955) and specificity (0.970; 95% CI: 0.963–0.975) values, as well as Macey et al., who found exfoliative cytology to outperform other diagnostic tests with sensitivity and specificity values of 0.900 (95% CI: 0.820 to 0.940) and 0.940 (95% CI: 0.880 to 0.970), respectively [20, 21].

The high DOR of 137.502 indicates that exfoliative cytology has good discriminative capacity. The PLR of 0.119 (95% CI: 0.090–0.159) suggests that cytology has a high capacity to confirm the presence of disease in truly affected individuals. Conversely, the negative likelihood ratio of 0.096 (95% CI: 0.059–0.158) indicates that there remains a 9.6% probability of disease even with a negative cytology result, highlighting that while the technique is valuable, it does not eliminate the need for confirmatory testing in suspicious cases.

While traditional exfoliative cytology using metal spatulas, wooden tongue depressors, and cotton-tipped applicators has historically been associated with poor sensitivity due to inadequate sampling of deeper epithelial layers, the introduction of cytobrushes has significantly revitalized oral cytology applications [8]. The most used brush types identified in our review include the cytobrush, OralCDx brush (OralScan Laboratories Inc., Suffern, Nueva York), and baby toothbrushes in resource-limited settings [22]. The OralCDx brush, which obtains transepithelial samples for Papanicolaou smears that can be analysed through image analysis systems, has shown varying but predominantly high sensitivity and specificity across studies [23].

LBC represents a significant advancement over conventional cytology, offering improved sample preservation and cellular morphological observation. Our analysis of studies utilizing LBC processing demonstrated generally higher sensitivity and specificity compared to conventional methods [24]. Notably, Navone et al. reported sensitivity and specificity of 95.1% and 99.0% using LBC compared to 85.7% and 95.9% using conventional exfoliative cytology in the same study. This improvement can be attributed to LBC's ability to produce more homogeneous samples with fewer air-drying artifacts and less obscuring elements such as blood, inflammation, mucus, and necrotic debris [21, 25].

The combination of cytological methods with molecular analyses has emerged as a promising approach to enhance diagnostic accuracy. Several studies in our review demonstrated significant improvements in sensitivity and specificity when cytology was combined with DNA-image cytometry, AgNOR analysis, or other molecular markers. For instance, Maraki et al. reported increased sensitivity and specificity of 100% and 97.4%, respectively, when combining cytological diagnosis with DNA cytometry [26]. Similarly, Kaur et al. showed an increase in sensitivity and specificity to 92% and 100%, respectively, after combining cytology with DNA cytometry. These findings suggest that the integration of molecular biomarkers with cytological methods could substantially improve the detection of oral malignancies and potentially malignant disorders [27].

Our meta-analysis faced several methodological challenges that merit discussion. To address the substantial heterogeneity observed across studies, we initially employed the random-effects model of DerSimonian and Laird to incorporate heterogeneity into the overall estimates [28]. However, recognising that these conventional methods might introduce bias in meta-analyses of binary outcomes [29], such as sensitivity and specificity, and that the normality assumption of estimates and its variance might not hold when dealing with few studies or sparse data [30], we

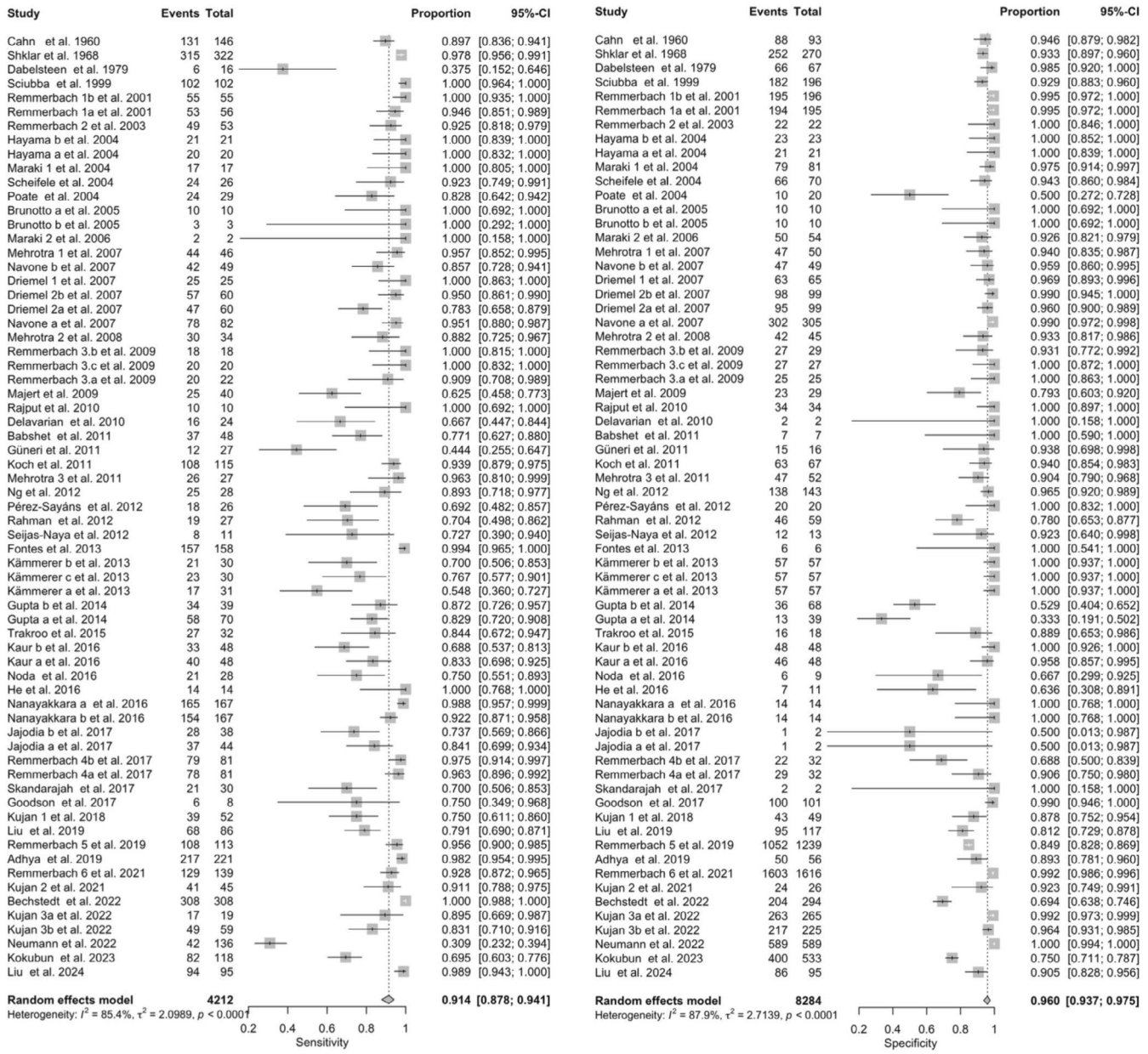

FIGURE 2 | Forest plot. The squares represent study-specific estimations, and their sizes correspond to relative weights. Diamonds represent pooled estimations and 95% CIs.

subsequently used bivariate mixed effects modelling for more reliable parameter estimates [31]. This approach explicitly accounts for the correlation between sensitivity and specificity, addressing the potential interdependence of these measures within individual studies and providing more robust summary estimates, particularly in the context of diagnostic test accuracy studies.

We conducted subgroup analyses to assess the robustness of our findings, systematically excluding individual studies to evaluate their influence on the pooled estimates. This approach allowed us to identify potential outliers and determine the stability of our results. These techniques aimed to explore potential sources of heterogeneity, including differences in study design, patient characteristics, cytological techniques, and reference standards [32].

Several potential sources of epidemiological bias must be considered when interpreting our results. Selection bias represents a significant concern, as many included studies employed convenience sampling rather than consecutive recruitment, potentially leading to overrepresentation of obvious or advanced cases (i.e., Will Rogers phenomenon) [33]. Studies varied considerably in their inclusion criteria, with some focusing on patients with clinical suspicion of malignancy while others included patients with and without symptoms, resulting in different pre-test probabilities. This spectrum bias may have influenced the apparent performance of cytological methods, as diagnostic tests often perform better in populations with higher disease prevalence or more advanced disease [34].

Additionally, verification bias may have affected our estimates, as not all patients in all studies underwent the reference standard

TABLE 2 | Subgroup analysis of cytology testing for oral squamous cell carcinoma and potentially malignant disorders diagnosis based on different covariates.

	No. of studies	DOR (95% CI)	I^2 (%)	*p*	Sensitivity (95% CI)	Specificity (95% CI)	PRL (95% CI)	NRL (95% CI)
Total	53	137.502 (79.733–237.127)	86.26	< 0.0001	0.914 (0.878–0.940)	0.960 (0.938–0.974)	11.970 (9.005–15.912)	0.096 (0.059–0.158)
Continent								
America	11	391.806 (216.089–710.412)	7	0.377	0.966 (0.930–0.984)	0.927 (0.892–0.952)	12.607 (7.978–19.922)	0.030 (0.014–0.066)
Europe	29	248.155 (114.957–535.689)	79.11	< 0.001	0.887 (0.814–0.934)	0.957 (0.928–0.975)	18.88 (11.954–29.819)	0.083 (0.038–0.180)
Asia	22	33.761 (16.061–70.966)	81.51	< 0.001	0.833 (0.772–0.880)	0.842 (0.769–0.896)	4.896 (3.354–7.147)	0.189 (0.129–0.279)
Oceania	5	134.774 (37.003–490.881)	72.55	0.006	0.826 (0.748–0.883)	0.965 (0.907–0.987)	23.555 (8.427–65.838)	0.187 (0.120–0.291)
Type of study								
Prospective	41	130.357 (71.371–238.092)	85.54	< 0.001	0.877 (0.840–0.907)	0.929 (0.901–0.950)	11.685 (8.505–16.053)	0.099 (0.052–0.191)
Retrospective	12	175. 982 (46.246–669.669)	88.36	< 0.001	0.899 (0.766–0.961)	0.940 (0.865–0.975)	14.269 (6.646–30.636)	0.093 (0.049–0.176)
Technique								
Conventional cytology	36	95.566 (47.932–190.537)	83.17	< 0.001	0.874 (0.826–0.910)	0.913 (0.869–0.943)	9.344 (6.403–13.638)	0.109 (0.046–0.259)
Liquid-based cytology	14	131.956 (35.610–488.971)	93.55	< 0.001	0.859 (0.742–0.928)	0.947 (0.898–0.973)	15.149 (7.958–28.84)	0.126 (0.062–0.256)
Cytology with complementary techniques	17	318.423 (116.529–870.109)	61.34	< 0.001	0.919 (0.848–0.958)	0.944 (0.903–0.969)	16.196 (9.016–29.093)	0.052 (0.028–0.097)
Instrument								
Cytobrush Plus GT	18	431.399 (177.839–1046.484)	64.06	< 0.001	0.897 (0.819–0.944)	0.973 (0.927–0.990)	3.1184 (12.112–80.284)	0.068 (0.021–0.222)
Cytobrush	15	89.937 (33.616–240.618)	67.23	< 0.001	0.876 (0.805–0.923)	0.887 (0.819–0.931)	7.056 (4.329–11.502)	0.112 (0.062–0.203)

(Continues)

TABLE 2 | (Continued)

	No. of studies	DOR (95% CI)	I^2 (%)	*p*	Sensitivity (95% CI)	Specificity (95% CI)	PRL (95% CI)	NRL (95% CI)
Orcellex brush	8	102.168 (25.658–406.831)	91.35	< 0.001	0.798 (0.603–0.911)	0.939 (0.888–0.968)	13.396 (7.006–25.616)	0.177 (0.064–0.491)
Toothbrush	12	80.013 (29.059–22.0314)	72.22	< 0.001	0.883 (0.815–0.928)	0.885 (0.765–0.948)	7.429 (3.731–14.792)	0.103 (0.046–0.227)
OralCDx	7	46.113 (8.960–237.329)	81.18	< 0.001	0.839 (0.689–0.925)	0.864 (0.730–0.937)	6.307 (2.627–15.142)	0.111 (0.022–0.576)
Tongue depressor	3	61.535 (2.040–1856.397)	97.43	< 0.001	0.924 (0.789–0.975)	0.832 (0.338–0.980)	6.587 (0.941–46.121)	0.110 (0.012–1019)
Metal spatula	2	1248.316 (273.627–5694.949)	12.99	0.284	0.928 (0.889–0.955)	0.988 (0.967–0.996)	80.522 (2.8417–228.165)	0.053 (0.019–0.151)
Sample size								
> 100	36	182.067 (85.850–386.122)	91.5	< 0.001	0.909 (0.862–0.941)	0.931 (0.899–0.953)	12.803 (8.881–18.457)	0.080 (0.040–0.158)
< 100	31	87.032 (43.760–173.094)	56.24	< 0.001	0.814 (0.757–0.860)	0.924 (0.891–0.948)	10.339 (6.849–15.609)	0.120 (0.078–0.184)
Diagnosed lesion								
OSCC	16	171.707 (68.807–428.494)	78.72	< 0.001	0.911 (0.855–0.947)	0.926 (0.870–0.958)	11.935 (6.817–20.898)	0.073 (0.038–0.141)
OPMDs	5	84.360 (28.446–250.180)	0	0.534	0.724 (0.405–0.910)	0.966 (0.904–0.988)	18.443 (7.736–43.972)	0.205 (0.086–0.488)
Both	47	132.768 (66.627–264.566)	88.17	< 0.001	0.880 (0.833–0.916)	0.928 (0.897–0.950)	11.563 (8.202–16.301)	0.098 (0.052–0.183)
QUADAS-2								
0–3 points	4	139.727 (12.487–1563.488)	55.48	0.081	0.803 (0.449–0.953)	0.973 (0.910–0.992)	20.816 (6.078–71.294)	0.125 (0.022–0.717)
4–5 points	27	57.957 (27.765–120.980)	83.99	< 0.001	0.851 (0.797–0.893)	0.871 (0.818–0.910)	6.492 (4.431–9.511)	0.144 (0.086–0.241)
6–7 points	36	237.397 (127.186–443.110)	77.05	< 0.001	0.909 (0.856–0.943)	0.951 (0.920–0.970)	17.832 (11.264–28.231)	0.072 (0.035–0.151)

Abbreviations: CI, Confidence Interval; DOR, Diagnostic Odds Ratio; NLR, Negative Likelihood Ratio; OPMDs, Oral Potentially Malignant Disorders); OSCC, Oral Squamous Cell Carcinoma; PLR, Positive Likelihood Ratio.

(histopathological examination) for non-suspicious lesions. In several studies, only patients with positive or suspicious cytology results received confirmatory biopsies, potentially inflating sensitivity estimates [7]. Furthermore, the lack of blinding between index test and reference standard interpretations in some studies may have introduced review bias [35].

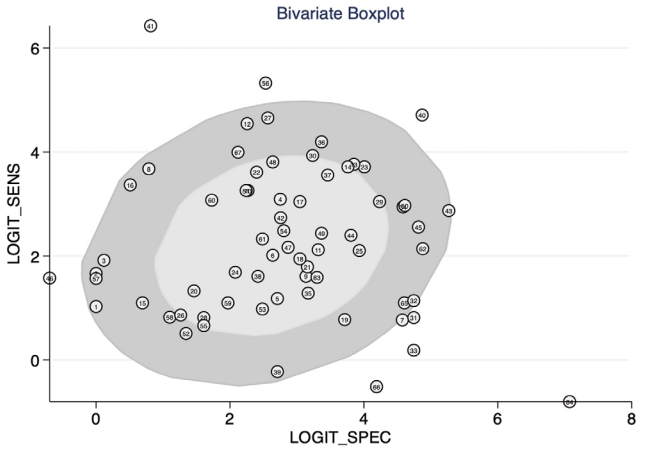

FIGURE 3 | Bivariate Boxplot. The bivariate boxplot represents the logit of sensitivity and specificity from the 67 statistical units, assessing the diagnostic capability of cytology for oral cancer and oral potentially malignant disorders.

Several shortcomings must be considered when interpreting our findings. First, the heterogeneity among included studies was substantial, influenced by factors such as varying sampling techniques, processing methods, and interpretation criteria. Second, there is a notable absence of a validated cytological classification system specifically designed for the oral cavity, which hampers standardization across studies. Third, most studies lacked long-term follow-up data on malignant transformation of OPMDs, limiting our ability to assess the prognostic value of cytological methods. Additionally, the geographical distribution of studies was uneven, with a paucity of research from African regions, potentially limiting the global generalizability of our findings.

Future directions for research should focus on establishing standardized protocols for sample collection, processing, and analysis to reduce procedural variability. The development and validation of a classification system specifically tailored for oral cytological abnormalities would enhance diagnostic reliability. Multicenter, longitudinal studies with proper sampling methods that reduce selection bias are needed to evaluate the utility of oral brush cytology in determining the malignant transformation risk of OPMDs. These studies should ensure proper blinding between index and reference tests and include both symptomatic and asymptomatic patients to better understand how test performance varies across different clinical scenarios. Additionally, research should continue exploring the integration

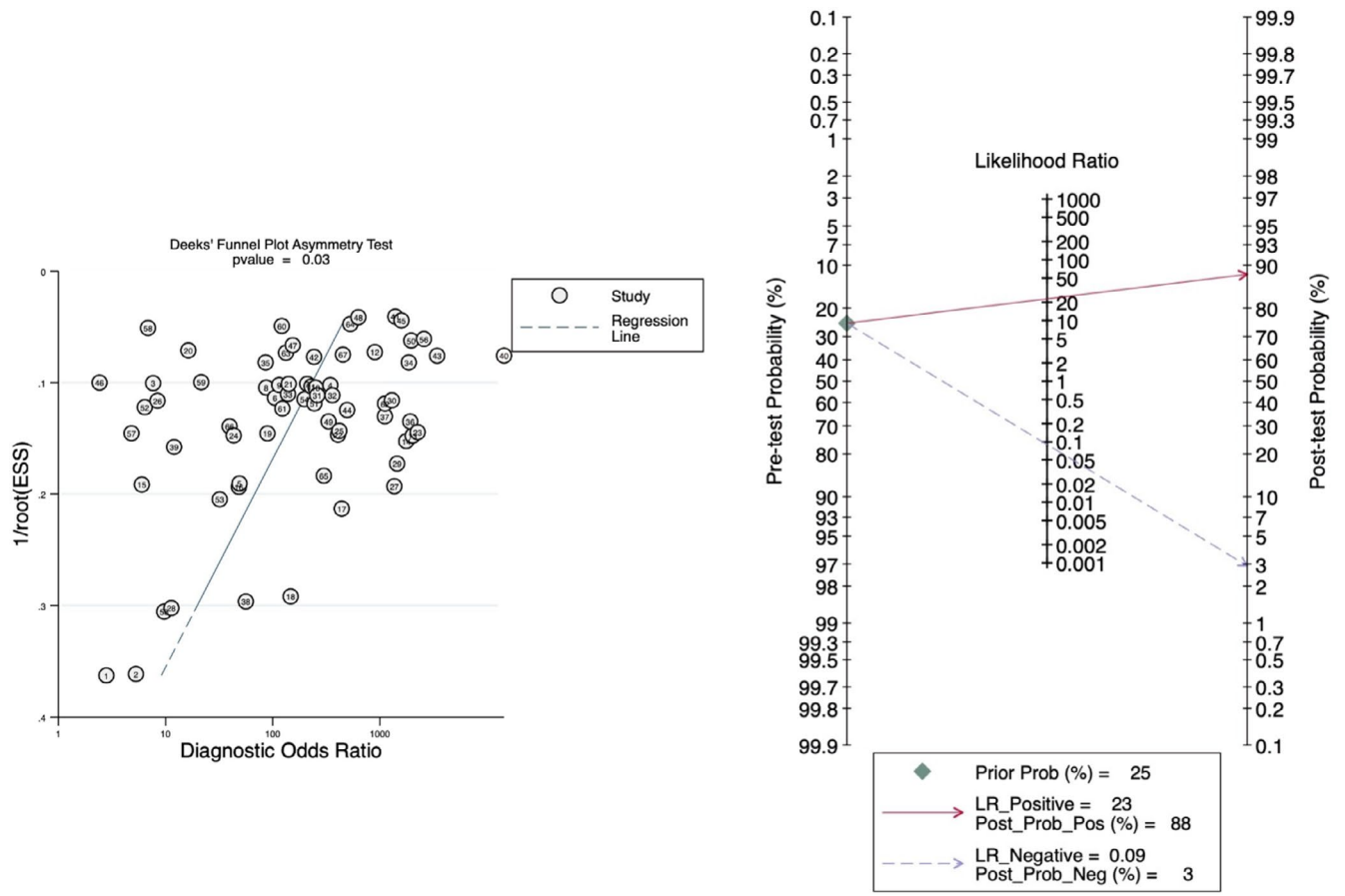

FIGURE 4 | (A) Funnel plot. The log estimates are represented on the x-axis, while their standard error (SE) values are shown on the y-axis. The bubbles represent primary-level- studies. (B) Fagan plot. This plot illustrates the post-test probabilities based on the pre-test probabilities and likelihood ratios, demonstrating the clinical utility of cytology for diagnosing oral squamous cell carcinoma and potentially malignant disorders.

of cytology with advanced molecular techniques to further improve diagnostic accuracy [7].

A more refined diagnostic approach could involve a two-step strategy to enhance efficiency while maintaining high sensitivity and specificity. Instead of subjecting all patients to confirmatory biopsies following cytology, a stratified diagnostic pathway could be considered. Patients with positive cytology results would proceed directly to histopathological confirmation, while those with negative results but high clinical suspicion could undergo adjunctive molecular testing or repeat cytology at a short interval. This strategy would optimize resource allocation, reducing unnecessary biopsies while maintaining diagnostic accuracy [35].

The potential impact of such an approach on patient adherence and overall diagnostic yield remains an open question. Previous studies have demonstrated that less invasive techniques such as brush cytology are more acceptable to patients compared to scalpel biopsy, potentially leading to higher participation rates in screening programs [36]. Similarly, studies in other fields have shown that sequential testing strategies can achieve comparable sensitivity while reducing unnecessary procedures [37]. Finally, the cost-effectiveness of implementing such a strategy in routine clinical practice would depend on multiple factors, including the prevalence of OSCC and OPMDs in the target population, adherence to follow-up recommendations, and the cutoff criteria for referral to biopsy.

5 | Conclusions

Oral brush cytology employing liquid-based technology offers advantages over conventional methods; however, its effectiveness is limited by variability in sampling techniques and diagnostic accuracy. While certain studies report poor sensitivity and specificity, others demonstrate promising results in OPMD and OSCC detection. Standardized cytological criteria, improved cyto-histopathological correlation, and longitudinal studies are essential to establish oral cytology as a reliable, minimally invasive diagnostic tool. Notwithstanding recent advances, the evidence compiled in this study reaffirms that histopathological examination remains the gold standard for the diagnosis and monitoring of these conditions.

Future research should prioritize the integrating cytological methods with molecular and imaging technologies to further improve diagnostic accuracy. Pragmatic trials assessing the effectiveness of brush cytology in real-world screening programs could provide valuable insights into its clinical utility and cost-effectiveness, ultimately guiding its incorporation into standardized protocols for early oral cancer detection.

Author Contributions

Hoda Tayebi-Hillali: data collection, data analysis, manuscript drafting. **Alejandro I. Lorenzo-Pouso:** conceptualization, methodology, supervision, manuscript review. **Xabier Marichalar-Mendía:** statistical analysis, data interpretation. **Pilar Gándara-Vila:** literature search, data collection, manuscript review. **Dolores Reboiras-López:** data collection, data analysis, manuscript review. **Andrés Blanco-Carrión:** supervision, manuscript drafting, final approval. **Martina Coppini:** literature search, data collection, manuscript review. **Vito Carlo Alberto Caponio:** project administration, manuscript drafting, final approval. **Mario Pérez-Sayáns:** project administration, manuscript drafting, final approval.

Acknowledgements

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that supports the findings of this study are available in the supplementary material of this article.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/jop.70010.

References

1. S. Warnakulasuriya, "Global Epidemiology of Oral and Oropharyngeal Cancer," *Oral Oncology* 45, no. 4-5 (2009): 309–316.

2. F. Bray, M. Laversanne, H. Sung, et al., "Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: a Cancer Journal for Clinicians* 74, no. 3 (2024): 229–263, https://doi.org/10.3322/caac.21834.

3. G. Sarode, N. Maniyar, S. C. Sarode, M. Jafer, S. Patil, and K. H. Awan, "Epidemiologic Aspects of Oral Cancer," *Disease-a-Month* 66, no. 12 (2020): 100988.

4. P. Güneri and J. B. Epstein, "Late Stage Diagnosis of Oral Cancer: Components and Possible Solutions," *Oral Oncology* 50, no. 12 (2014): 1131–1136.

5. S. Warnakulasuriya, O. Kujan, J. M. Aguirre-Urizar, et al., "Oral Potentially Malignant Disorders: A Consensus Report From an International Seminar on Nomenclature and Classification, Convened by the WHO Collaborating Centre for Oral Cancer," *Oral Diseases* 27, no. 8 (2021): 1862–1880, https://doi.org/10.1111/odi.13704.

6. O. Iocca, T. P. Sollecito, F. Alawi, et al., "Potentially Malignant Disorders of the Oral Cavity and Oral Dysplasia: A Systematic Review and Meta-Analysis of Malignant Transformation Rate by Subtype," *Head & Neck* 42, no. 3 (2020): 539–555.

7. M. Diniz-Freitas, A. García-García, A. Crespo-Abelleira, J. L. Martins-Carneiro, and J. M. Gándara-Rey, "Applications of Exfoliative Cytology in the Diagnosis of Oral Cancer," *Medicina Oral* 9, no. 4 (2004): 355–361 English, Spanish.

8. M. Pérez-Sayáns, J. M. Somoza-Martín, F. Barros-Angueira, et al., "Exfoliative Cytology for Diagnosing Oral Cancer," *Biotechnic & Histochemistry* 85, no. 3 (2010): 177–187, https://doi.org/10.3109/10520290903162730.

9. V. C. A. Caponio, F. F. E. Silva, F. Popolo, et al., "State of Art of Micronuclei Assay in Exfoliative Cytology as a Clinical Biomarker of Genetic Damage in Oral Carcinogenesis: A Systematic Review and Meta-Analysis," *Mutation Research, Reviews in Mutation Research* 794 (2024): 108508, https://doi.org/10.1016/j.mrrev.2024.108508.

10. A. H. Alsarraf, O. Kujan, and C. S. Farah, "The Utility of Oral Brush Cytology in the Early Detection of Oral Cancer and Oral Potentially Malignant Disorders: A Systematic Review," *Journal of Oral Pathology & Medicine* 47, no. 2 (2018): 104–116, https://doi.org/10.1111/jop.12660.

11. P. Brocklehurst, O. Kujan, L. A. O'Malley, G. Ogden, S. Shepherd, and A. M. Glenny, "Screening Programmes for the Early Detection

and Prevention of Oral Cancer," *Cochrane Database of Systematic Reviews* 2013, no. 11 (2013): CD004150, https://doi.org/10.1002/14651858.CD004150.pub4.

12. M. D. F. McInnes, D. Moher, B. D. Thombs, et al., "Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement," *JAMA* 319, no. 4 (2018): 388–396, https://doi.org/10.1001/jama.2017.19163.

13. A. B. Olawaiye, T. P. Baker, M. K. Washington, and D. G. Mutch, "The New (Version 9) American Joint Committee on Cancer Tumor, Node, Metastasis Staging for Cervical Cancer," *CA: a Cancer Journal for Clinicians* 71, no. 4 (2021): 287–298, https://doi.org/10.3322/caac.21663.

14. F. F. V. E. Silva, V. C. A. Caponio, M. Pérez-Sayáns, et al., "Tumor Budding Is a Prognostic Factor in Head and Neck Squamous Cell Carcinoma: A Comprehensive Meta-Analysis and Trial Sequential Analysis," *Critical Reviews in Oncology/Hematology* 193 (2024): 104202, https://doi.org/10.1016/j.critrevonc.2023.104202.

15. Q. X. Huang and X. W. Huang, "QUADAS-2 Tool for Quality Assessment in Diagnostic Meta-Analysis," *Annals of Palliative Medicine* 11, no. 5 (2022): 1844–1845, https://doi.org/10.21037/apm-22-204.

16. P. F. Whiting, A. W. Rutjes, M. E. Westwood, et al., "QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies," *Annals of Internal Medicine* 155, no. 8 (2011): 529–536.

17. J. A. Swets, "Measuring the Accuracy of Diagnostic Systems," *Science* 240, no. 4857 (1988): 1285–1293, https://doi.org/10.1126/science.3287615.

18. J. P. T. Higgins, S. G. Thompson, J. J. Deeks, and D. G. Altman, "Measuring Inconsistency in Meta-Analyses," *BMJ (Clinical Research Ed.)* 327 (2003): 557, https://doi.org/10.1136/bmj.327.7414.557.

19. M. Egger, G. D. Smith, M. Schneider, and C. Minder, "Bias in Meta-Analysis Detected by a Simple, Graphical Test," *BMJ (Clinical Research Ed.)* 315, no. 7109 (1997): 629, https://doi.org/10.1136/bmj.315.7109.629.

20. S. Dolens Eda, F. V. Nakai, J. L. Santos Parizi, and G. Alborghetti Nai, "Cytopathology: A Useful Technique for Diagnosing Oral Lesions?: A Systematic Literature Review," *Diagnostic Cytopathology* 41, no. 6 (2013): 505–514.

21. R. Macey, T. Walsh, P. Brocklehurst, et al., "Diagnostic Tests for Oral Cancer and Potentially Malignant Disorders in Patients Presenting With Clinically Evident Lesions," *Cochrane Database of Systematic Reviews* 29, no. 5 (2015): CD010276.

22. C. Scheifele, A. M. Schmidt-Westhausen, T. Dietrich, and P. A. Reichart, "The Sensitivity and Specificity of the OralCDx Technique: Evaluation of 103 Cases," *Oral Oncology* 40, no. 8 (2004): 824–828, https://doi.org/10.1016/j.oraloncology.2004.02.004.

23. J. J. Sciubba and U.S. Collaborative OralCDx Study Group, "Improving Detection of Precancerous and Cancerous Oral Lesions. Computer-Assisted Analysis of the Oral Brush Biopsy," *Journal of the American Dental Association* 130, no. 10 (1999): 1445–1457, https://doi.org/10.14219/jada.archive.1999.0055.

24. O. Kujan, M. Idrees, N. Anand, B. Soh, E. Wong, and C. S. Farah, "Efficacy of Oral Brush Cytology Cell Block Immunocytochemistry in the Diagnosis of Oral Leukoplakia and Oral Squamous Cell Carcinoma," *Journal of Oral Pathology & Medicine* 50, no. 5 (2021): 451–458, https://doi.org/10.1111/jop.13153.

25. R. Navone, P. Burlo, A. Pich, et al., "The Impact of Liquid-Based Oral Cytology on the Diagnosis of Oral Squamous Dysplasia and Carcinoma," *Cytopathology* 18, no. 6 (2007): 356–360, https://doi.org/10.1111/j.1365-2303.2006.00402.x.

26. D. Maraki, J. Becker, and A. Boecking, "Cytologic and DNA-Cytometric Very Early Diagnosis of Oral Cancer," *Journal of Oral Pathology & Medicine* 33, no. 7 (2004): 398–404, https://doi.org/10.1111/j.1600-0714.2004.0235.x.

27. M. Kaur, U. Handa, H. Mohan, and A. Dass, "Evaluation of Brush Cytology and DNA Image Cytometry for the Detection of Cancer of the Oral Cavity," *Diagnostic Cytopathology* 44, no. 3 (2016): 201–205, https://doi.org/10.1002/dc.23422.

28. R. DerSimonian and N. Laird, "Meta-Analysis in Clinical Trials Revisited," *Contemporary Clinical Trials* 45 (2015): 139–145, https://doi.org/10.1016/j.cct.2015.09.002.

29. J. L. Tang, "Weighting Bias in Meta-Analysis of Binary Outcomes," *Journal of Clinical Epidemiology* 53, no. 11 (2000): 1130–1136, https://doi.org/10.1016/s0895-4356(00)00237-7.

30. Y. Takwoingi, B. Guo, R. D. Riley, and J. J. Deeks, "Performance of Methods for Meta-Analysis of Diagnostic Test Accuracy With Few Studies or Sparse Data," *Statistical Methods in Medical Research* 26, no. 4 (2017): 1896–1911, https://doi.org/10.1177/0962280215592269.

31. H. Chu, L. Nie, Y. Chen, Y. Huang, and W. Sun, "Bivariate Random Effects Models for Meta-Analysis of Comparative Studies With Binary Outcomes: Methods for the Absolute Risk Difference and Relative Risk," *Statistical Methods in Medical Research* 21, no. 6 (2012): 621–633, https://doi.org/10.1177/0962280210393712.

32. A. D. Oxman and G. H. Guyatt, "A Consumer's Guide to Subgroup Analyses," *Annals of Internal Medicine* 116, no. 1 (1992): 78–84, https://doi.org/10.7326/0003-4819-116-1-78.

33. A. R. Feinstein, D. M. Sosin, and C. K. Wells, "The Will Rogers Phenomenon. Stage Migration and New Diagnostic Techniques as a Source of Misleading Statistics for Survival in Cancer," *New England Journal of Medicine* 312, no. 25 (1985): 1604–1608, https://doi.org/10.1056/NEJM198506203122504.

34. M. Delgado-Rodríguez and J. Llorca, "Bias," *Journal of Epidemiology and Community Health* 58, no. 8 (2004): 635–641, https://doi.org/10.1136/jech.2003.008466.

35. R. Neugebauer and S. Ng, "Differential Recall as a Source of Bias in Epidemiologic Research," *Journal of Clinical Epidemiology* 43, no. 12 (1990): 1337–1341, https://doi.org/10.1016/0895-4356(90)90100-4.

36. C. L. Lavelle and C. Scully, "Criteria to Rationalize Population Screening to Control Oral Cancer," *Oral Oncology* 41, no. 1 (2005): 11–16, https://doi.org/10.1016/j.oraloncology.2004.04.009.

37. P. Dwivedi, A. Lohiya, P. Bahuguna, et al., "Cost-Effectiveness of Population-Based Screening for Oral Cancer in India: An Economic Modelling Study," *Lancet Regional Health - Southeast Asia* 16 (2023): 100224, https://doi.org/10.1016/j.lansea.2023.100224.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.