

Research Article

Salivary Exosomal MicroRNAs for Detection of Oral Squamous Cell Carcinoma: A Systematic Review

Arindam Chakraborty¹, Arun Pandey² and Ankit Kumar Tiwari^{3*}

¹Associate Professor, Department of Microbiology, Autonomous State Medical College, Kaushambi Uttar Pradesh, India

²Assistant Professor, Department of Dentistry, Autonomous State Medical College, Kaushambi Uttar Pradesh, India

³Department of Biochemistry, Autonomous State Medical College, Kaushambi Uttar Pradesh, India

OPEN ACCESS

Corresponding Author:

Dr. Ankit Kumar Tiwari
(biochemistankit@gmail.com)

Received: 20 June 2026
Revised: 12 July 2026
Accepted: 5 August 2026
Published: 20 September 2026

Doi: 10.61336/ML.20263002.05

©2026 the Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)

ABSTRACT

Background: Salivary extracellular vesicle-associated microRNAs offer a minimally invasive means of investigating oral squamous cell carcinoma (OSCC). Their clinical value depends on discrimination within relevant patient populations, reproducible laboratory procedures and validation of thresholds. Evidence for recognizing established cancer must be distinguished from evidence for detecting early-stage disease. **Objective:** To evaluate human diagnostic evidence for salivary exosomal microRNAs in OSCC, examine methodological limitations and determine whether published results support early detection or quantitative pooling. **Methods:** A restricted systematic search of Europe PMC covered publications first published from 1 January 2015 through 31 March 2025. Targeted publisher and bibliographic retrieval supplemented the search. Human studies reporting salivary extracellular vesicle microRNA diagnostic performance were considered; mixed oral and oropharyngeal cohorts were analyzed separately. Available reports were appraised using a provisional QUADAS-2 framework. Heterogeneity, reporting discrepancies and unavailable diagnostic contingency data precluded meta-analysis. **Results:** The broad search returned 226 records. Eleven were retained for detailed source assessment and one additional candidate was identified through citation checking. Six diagnostic reports were retained: four OSCC-focused and two mixed-site reports. Five were assessed in full text and one from its indexed abstract. Reported AUCs varied from 0.59 to 0.99 across different markers and populations. Near-perfect performance in small developmental cohorts coexisted with substantially lower discrimination in other studies. Early-stage subgroup evidence was limited and no retained report established effectiveness in population screening. Selection bias, data-dependent thresholds, uncertain participant independence and restricted validation weakened applicability. **Conclusion:** Salivary exosomal microRNAs remain investigational biomarkers. The available evidence supports further validation but does not establish a pooled diagnostic accuracy, a superior clinical panel or a safe rule-out test for early OSCC. A complete multi-database review and prospective diagnostic cohorts are required before clinical adoption.

Keywords: Oral Squamous Cell Carcinoma, Saliva, Extracellular Vesicles, Exosomes, MicroRNAs, Diagnostic Accuracy, Early Detection, Systematic Review

INTRODUCTION

Clinical Importance and the Intended Diagnostic Problem

Oral cancer contributes to the global cancer burden, but estimates for lip and oral cavity cancer should not be treated as estimates for every histological subtype or for the entire head-and-neck region. International cancer statistics provide the epidemiological setting for improved detection while preserving these anatomical distinctions [1]. For OSCC, an effective diagnostic adjunct would identify patients needing definitive investigation at a clinically useful point in their care. Demonstrating a molecular difference between patients already known to have cancer and healthy volunteers is an initial step, rather than proof that an assay improves diagnosis in practice.

The meaning of “early” requires particular care. A small primary tumor, a stage I or II cancer, a lesion detected before symptoms and a cancer identified sooner after presentation represent different outcomes. Oral cavity staging incorporates factors that cannot be inferred from salivary expression alone. Oropharyngeal cancer also requires separate

consideration because site and HPV-associated staging affect case mix and prognosis [2]. Consequently, a mixed head-and-neck cohort cannot automatically supply an accuracy estimate for early OSCC. The clinically relevant question is whether an assay distinguishes early OSCC from the lesions and inflammatory conditions encountered where testing would occur.

Biological and Analytical Basis

Saliva contains material originating from oral tissues, salivary glands and other sources. Earlier research demonstrated the feasibility of measuring salivary microRNAs for oral cancer detection, including reduced miR-200a and miR-125a in cancer samples [3]. However, whole saliva and isolated vesicle fractions are different analytical specimens. Findings from one compartment cannot establish the direction, concentration or diagnostic threshold of a marker in another. The distinction is essential when interpreting the supplied abstract’s statements about particular microRNAs.

Extracellular vesicles are membrane-delimited particles that carry molecular cargo. The term “exosome” implies a particular biogenesis that routine size measurement alone does not establish. MISEV2018 and MISEV2023 recommend descriptions that reflect the demonstrated characteristics of the preparation [4,5]. This manuscript retains “exosomal” when describing the terminology of primary studies, but uses “extracellular vesicle-associated” where appropriate. A positive vesicle marker, a small particle diameter and a precipitation pellet do not independently prove tumor origin or a pure exosome population.

Vesicle-associated microRNAs could be useful because the carrier may protect RNA and because tumor-associated alterations may be measurable in saliva. Nevertheless, a stable molecule is not necessarily a disease-specific molecule. An assay must separate variation caused by cancer from variation caused by collection, oral inflammation, tobacco exposure and laboratory processing. The present review therefore evaluates diagnostic design and analytical validity together. It asks what the reported measurements demonstrate, which conclusions depend on untested assumptions and which additional studies would change clinical confidence.

METHODS

Review Design and Reporting Status

The manuscript follows the organization of PRISMA 2020 and the diagnostic test accuracy extension, PRISMA-DTA [6,7]. These guidelines informed reporting; complete compliance is not claimed. The supplied abstract was treated as a topic brief, not as an authenticated search log or dataset. No prospective registration, protocol identifier, independent reviewer pair or author correspondence is claimed. Eligibility decisions and source extraction were undertaken in one assisted workflow, with internal consistency checks rather than independent duplicate assessment.

A restricted systematic evidence verification was performed within the original publication window. The retrospective design, single searchable bibliographic interface and incomplete access to some full texts prevent this draft from being represented as a completed, comprehensive multi-database systematic review. These limitations are stated explicitly because journal-style formatting cannot substitute for review conduct. The documented results are suitable for checking and extending by the dissertation team; they should not be relabeled as results from searches that were never executed.

Review Question and Objectives

The Diagnostic Question Was: Among adults with suspected or established oral malignancy, how accurately do salivary extracellular vesicle-associated microRNAs distinguish histologically established OSCC from non-cancer comparators and what evidence specifically supports early-stage detection? A population, index test, comparator, outcomes and study design framework was used. Histopathology was the intended reference standard for cancer. Comparator groups included healthy individuals and clinically relevant non-cancer lesions, with their different applicability considered separately.

The primary objective was to describe verified diagnostic performance without conflating OSCC-specific results with mixed-site results. Secondary objectives were to examine early-stage representation, identify sources of clinical and analytical heterogeneity, assess risk of bias, evaluate whether diagnostic estimates were reproducible and define requirements for future validation. Prognostic associations and mechanistic experiments were contextual evidence only unless the report also contained an eligible human salivary diagnostic comparison (Table 1).

Table 1: Diagnostic Review Framework

Element	Operational definition
Population	Adults with OSCC; mixed oral and oropharyngeal cohorts analyzed as indirect evidence.
Index test	A microRNA or panel measured in a salivary extracellular vesicle preparation.
Comparator	Healthy controls or non-cancer lesions; distinguish comparator spectra.
Reference standard	Histopathology for cancer; document how non-cancer status is established.
Outcomes	AUC, sensitivity and specificity; early-stage performance, external validation and clinical utility.
Design	Human diagnostic comparisons with eligible numerical accuracy reporting.

Eligibility Criteria

Eligible diagnostic reports were peer-reviewed human studies within the specified window that measured microRNAs in a salivary vesicle preparation, included cancer and non-cancer samples and reported an AUC or a sensitivity-specificity result. OSCC-focused studies formed the direct evidence stratum. Mixed oral and oropharyngeal studies were retained in an explicitly indirect stratum because excluding them from all discussion would obscure the origin of several frequently cited biomarkers. Their combined estimates were not treated as oral-cavity-specific estimates.

Reviews, preprints, conference records without a qualifying report, animal studies, cell-line-only studies, tissue-only analyses and plasma-only assays were excluded from the diagnostic synthesis. Salivary studies of proteins, cytokines or other RNA classes were not eligible merely because vesicles were isolated. Expression differences, pathway enrichment and clustering without a reported diagnostic performance measure did not meet the diagnostic outcome criterion. Reports relevant to these exclusions were retained as contextual references where they explained a methodological or biological issue.

Information Sources and Actual Searches

Europe PMC was queried through its public application programming interface on 19 September 2026. An exploratory condition-specific query returned 36 records. Because it missed a known mixed-site study, a broader saliva-vesicle-microRNA query was executed and returned 226 records. The broader set formed the screening denominator; the exploratory set was not added again. The full query, date restriction and distinction between executed and proposed database strategies appear in Appendix A. This follows the transparency principle of PRISMA-S [8].

The retrieved records were screened by title and, for potentially relevant records, indexed abstract. Bibliographic references and targeted publisher searches were used to obtain primary sources and check candidate studies. One additional candidate, Farag and colleagues, entered detailed assessment through citation checking. Direct searches of PubMed's own interface, Scopus, Embase, Web of Science and CENTRAL were not performed. Europe PMC retrieval includes biomedical bibliographic records but is not described as an equivalent execution of those platforms.

Selection and Extraction

Records were treated as reports, not automatically as independent participant cohorts. Eleven candidates from the broad search were retained for detailed assessment. The additional citation-derived report brought the candidate set to twelve. Available full texts, indexed abstracts and publisher records were inspected for eligibility and source-level outcomes. Six diagnostic reports were retained, including one with abstract-only extraction. Formal full-text exclusion decisions remain provisional for candidates whose complete articles could not be obtained.

Extraction captured publication identity, country, recruitment design, oral site, sample source, discovery and validation sets, control selection, assay method, reported AUCs, available uncertainty measures and restrictions affecting early detection. Tissue samples, pooled controls and external genomic datasets were not added to salivary patient totals. Missing values were recorded as unavailable rather than reconstructed from rounded percentages. The evidence tables distinguish publication-level sample descriptions from the denominators actually used in diagnostic testing.

Risk of Bias and Certainty Assessment

QUADAS-2 was selected because the target question concerns diagnostic accuracy [9]. The four domains were patient selection, index test, reference standard and flow and timing. Applicability was considered separately for population, index test and reference standard. Particular attention was paid to healthy-control designs, selection after diagnosis, threshold optimization and the use of the same samples for model development and evaluation. STARD 2015 informed checks for diagnostic reporting, but was not used as a numerical quality score [10].

Judgments in this document are provisional and explicitly supported by the available report. An unclear judgment denotes insufficient information; it does not assert that a procedure was performed incorrectly. Abstract-only evidence was not assigned reassuring low-risk judgments. Certainty was considered at the level of a clinical claim rather than averaged across a paper. The diagnostic GRADE framework guided consideration of bias, indirectness, inconsistency, imprecision and publication bias [11]. A formal consensus GRADE profile was not completed.

Synthesis and Conditions for Meta-Analysis

The primary synthesis was narrative, organized by diagnostic purpose and evidence directness. AUCs were displayed descriptively, without inverse-variance pooling or a pooled marker-family effect. If sufficient compatible studies become available, a bivariate random-effects model can jointly synthesize sensitivity and specificity while allowing their correlation [12]. That approach requires defensible diagnostic counts and clinical comparability; it does not justify combining unrelated microRNAs into a single test.

Meta-analysis was not conducted because marker identity, assay units, cutoffs, disease sites and validation designs differed and complete patient-level two-by-two tables were unavailable. Multiple markers measured in the same participants were not treated as independent studies. Funnel plots, heterogeneity statistics and summary ROC curves were not generated because the required evidence structure was absent. Figures show documented source results or the actual review workflow; none represents an invented pooled analysis.

PRISMA Style Identification and Selection Results

The reproducible broad search identified 226 records. No additional exact-record deduplication was applied within that single returned set; this does not establish independence of the underlying cohorts. After title and relevant abstract screening, 215 records did not proceed to detailed assessment and eleven did. Citation checking supplied one additional report. Among the twelve candidates, nine full texts and three abstract or bibliographic-level sources were available. Six reports met the diagnostic synthesis criteria, including five with full-text assessment and one with abstract-only assessment.

Six candidate reports were not included in the diagnostic synthesis: two presented human salivary expression or cargo profiling without an eligible numerical diagnostic accuracy endpoint; one was cell-line-only; one used computational data set rather than a qualifying patient salivary diagnostic cohort; one concerned analytical sampling feasibility; and one was an expression study without a verified qualifying accuracy estimate in the available source. These are restricted-audit counts, not a reconstruction of the supplied abstract’s unrecorded four-database search. Figure 1 and Table 2 preserve that distinction.

Study Characteristics

Six Reports Constitute the Retained Diagnostic Evidence Set: Langevin and colleagues, Gai and colleagues, He and colleagues, Faur and colleagues and two reports by Patel and colleagues [13-18]. Four focused on OSCC and two included oral and oropharyngeal disease. All investigated samples from diagnosed patients rather than demonstrating a prospective population-screening pathway. Discovery experiments, human validation samples and tissue-based corroboration were combined to different degrees, making a single aggregate participant count misleading (Table 3).

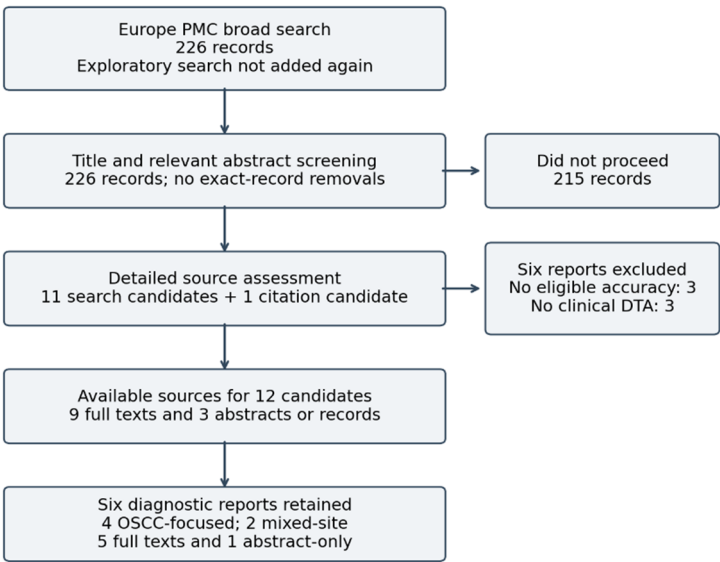


Figure 1: Restricted Search and Source-Assessment Flow.
This is An Evidence-Audit Diagram, not a Completed Four-Database PRISMA Flow. Candidate Exclusion Categories Include Provisional Abstract-Level Decisions

Table 2: Actual Restricted Search and Assessment Counts

Stage	Count	Interpretation
Broad search records	226	Europe PMC; executed query in Appendix A.
Exact-record removals	0	Single returned set; cohort overlap still possible.
Records screened	226	Title and relevant abstract assessment.
Did not proceed	215	Outside diagnostic candidate set.
Search candidates	11	Detailed source assessment.
Citation candidate	1	Farag report.
Total candidates	12	9 full texts; 3 abstract or record-level sources.
Not retained	6	Reasons in Appendix B.
Diagnostic reports	6	4 OSCC-focused and 2 mixed-site; not six proven independent cohorts.
Full-text included reports	5	One additional retained report was abstract-only.

The two Patel reports arose from related research settings and used similar discovery sample structures. Participant overlap could not be excluded from the published descriptions reviewed here. Accordingly, no unique combined sample total is reported. He's indexed abstract supplies a usable AUC and the validation comparison, but not the complete information needed for definitive risk-of-bias assessment. These restrictions are carried into the tables rather than concealed by a nominal total sample size (Table 4).

Review of Literature and Thematic Evidence Synthesis

Single-Marker OSCC Evidence: Gai's investigation used 21 OSCC cases and eleven controls across discovery and follow-up sample sets. Reported AUCs were 0.847 for miR-512-3p and 0.871 for miR-412-3p. Thresholds were selected using Youden's index [14]. These findings establish preliminary case-control discrimination. They do not establish a fixed clinical threshold because threshold selection and performance estimation were closely linked. The small control set also limits confidence about specificity in the broader range of non-malignant oral conditions.

He's abstract reports microarray discovery in four cases and four controls, followed by qRT-PCR evaluation of 45 preoperative OSCC cases and ten controls. The reported AUC for miR-24-3p was 0.738 [15]. This value directly contradicts an interpretation that all leading exosomal candidates routinely achieve AUCs above 0.85. Sensitivity and specificity values from secondary summaries were not imported into the extraction as though verified in the primary full text. A modest AUC may still motivate assay development, but the operating point and relevant false-negative rate remain essential for clinical interpretation.

Mixed-Site Evidence and Inconsistent Marker Behavior

Langevin's human ddPCR comparison involved eleven cancers and nine controls, following a separate five-case/five-control pilot. Only three validation cancers arose in the oral cavity. Reported sensitivity and specificity were 18% and 100% for miR-10b-5p and 45% and 89% for miR-486-5p [13]. These source-reported estimates are retained without reconstructing exact contingency tables because narrative concentrations and threshold descriptions contain inconsistencies. Identifying one early-stage case does not establish early-stage sensitivity.

Table 3: Characteristics of Retained Diagnostic Reports

Study and setting	Design and saliva samples	Index and comparator	Principal limitation
Langevin 2017 [13] USA	Case-control; pilot 5 cases/5 controls; separate ddPCR 11/9. Validation: 3 oral, 8 oropharyngeal.	Ultracentrifugation; sequencing then ddPCR; cancer-free controls.	Mixed sites; demographic imbalance; 1 early-stage validation cancer.
Gai 2018 [14] Italy	Case-control; discovery 5/5; additional 16 OSCC/6 controls. Total 21/11.	Charge-based EV precipitation; qRT-PCR; matched healthy controls.	Small control group; data-selected cutoffs.
He 2020 [15] China	Discovery 4/4; qRT-PCR 45 preoperative OSCC/10 controls. Overlap not resolved.	Salivary exosomal miR-24-3p; healthy controls; microarray then qRT-PCR.	Abstract-only extraction; detailed isolation and stage distribution unverified.
Faur 2022 [16] Romania	Case-control; recruited 25 cancers/25 controls; miRNA results describe 24 cancers.	Ultracentrifugation and qRT-PCR; miR-486-5p/miR-10b-5p; healthy volunteers.	Mixed sites; sample-flow and stage-II reporting discrepancies.
Patel 2022 [17] India	Discovery 8 cancer saliva samples and pooled controls; validation 12 cancer/5 healthy saliva samples.	Precipitation and qRT-PCR; miR-1307-5p; tissue work analyzed separately.	Selected male buccal-cancer population; unclear diagnostic unit; possible cohort overlap.
Patel 2023 [18] India	Discovery table: 8 patient saliva samples, pooled controls; validation 20 patient/5 healthy saliva samples.	Precipitation and qRT-PCR; miR-140-5p/143-5p/145-5p; non-cancer controls.	Sampling includes tissue/leukoplakia work; exact salivary ROC denominator needs clarification.

Table 4: Verified Diagnostic Results and Interpretation

Report and marker	Reported result	Interpretive restriction
Gai [14] miR-512-3p; miR-412-3p	AUC 0.847; 0.871	Not externally validated clinical cutoffs; correlated markers in shared samples.
He [15] miR-24-3p	AUC 0.738; P=0.02	Abstract-verified; no primary full-text sensitivity/specificity extraction.
Langevin [13] miR-10b-5p	Sensitivity 18%; specificity 100%	Source-reported, mixed cohort; no reconstructed 2x2 table.
Langevin [13] miR-486-5p	Sensitivity 45%; specificity 89%	Threshold narrative inconsistencies; no pooled estimate.
Faur [16] miR-486-5p	Overall AUC 0.72 (95% CI 0.57–0.88); oral-cavity AUC 0.67	Combined-site estimate does not equal early-OSCC accuracy.
Faur [16] miR-10b-5p	Overall AUC 0.59 (95% CI 0.43–0.76)	Weak discrimination; interval includes 0.50.
Patel 2022 [17] miR-1307-5p	AUC 0.99; sensitivity/specificity reported as 99.99% across sample types	Unclear empirical patient-level meaning; not a validated near-perfect test.
Patel 2023 [18] Three-miRNA panel	Salivary AUC 0.99; sensitivity 98%; specificity 99%	Not convertible into integer counts for the stated small validation set.

Faur's study recruited 25 cancer patients and 25 controls, with miRNA analyses describing 24 cancer samples. Overall AUCs were 0.72 for miR-486-5p and 0.59 for miR-10b-5p; the oral-cavity miR-486-5p subgroup AUC was 0.67 [16]. The report contains disagreement between text and table for a stage II AUC. Consequently, the disputed early-stage value is not plotted or used to support a screening claim. Directional variation in miR-10b-5p and weaker discrimination than initially anticipated illustrate why a candidate's name cannot stand in for a validated assay.

High-Performing Developmental Signatures

Patel's 2022 report evaluated miR-1307-5p and described a salivary validation set of twelve cancer samples and five healthy controls, alongside discovery, tissue and external genomic analyses. It reported an AUC of 0.99 and near-perfect sensitivity and specificity across sample types [17]. The reported percentages are not directly interpretable as empirical patient fractions in that small salivary set. They should therefore be regarded as developmental model results requiring clarification, rather than rounded counts suitable for meta-analysis.

The subsequent three-miRNA report evaluated miR-140-5p, miR-143-5p and miR-145-5p. Its table describes twenty patient saliva samples and five healthy saliva samples in validation. A salivary panel AUC of 0.99, sensitivity of 98% and specificity of 99% were reported [18]. Those percentages likewise cannot be converted into integer diagnostic counts for the stated validation set without additional information. Tissue confirmation and functional experiments strengthen biological motivation, but they do not independently validate patient-level salivary specificity or establish the panel's incremental value beyond examination.

Contextual Evidence and Boundaries of Inclusion

Two Hofmann reports investigated salivary exosome cargo and paired plasma-saliva miRNA profiles in HNSCC [19] [20]. Their molecular patterns are relevant to assay development, but the absence of an eligible numerical diagnostic accuracy endpoint in the assessed reports prevents their inclusion in the accuracy set. These studies also illustrate why plasma and salivary profiles require separate evaluation. A correlated molecular profile is not a substitute for a prespecified decision rule tested against an independent reference standard.

Coon's miR-365 work used oral cancer cell lines; Jayaseelan's TP53 investigation used computational resources; and Yap's oral-swirl study addressed collection and RNA protection [21-23]. Each contributes a different type of evidence, but none supplies a qualifying human OSCC diagnostic comparison. Farag's report was identified as a salivary miR-134/miR-200a expression study; complete primary-text assessment was unavailable and no qualifying accuracy estimate was verified [24]. It remains a candidate requiring final author assessment rather than a source of inferred sensitivity.

Non-Cancer Lesions and Biomarker Specificity

Byun's study compared oral lichen planus with healthy controls and identified altered salivary exosomal microRNAs, including miR-4484 [25]. It is not an OSCC diagnostic study, but it demonstrates the importance of testing disease controls. A microRNA associated with an inflammatory or potentially malignant condition may be useful for a separate question without being specific for invasive carcinoma. The supplied abstract's mention of miR-1246 should therefore not be interpreted as independently verified OSCC screening evidence from an oral lichen planus comparison.

No retained report established a transferable rule for distinguishing early OSCC from the full clinical spectrum of oral potentially malignant disorders. This gap matters because the relevant comparator in a referral clinic is often another lesion, not normal mucosa. A test developed against healthy volunteers may exploit biological contrasts that disappear when patients share inflammation, tobacco exposure or epithelial dysplasia. Comparative studies must therefore define the diagnostic target before selecting controls, rather than broadening the claimed indication after obtaining a favorable ROC curve.

Risk of Bias Assessment

The QUADAS-2 assessment identifies high concern about patient selection and applicability to early detection. Recruiting known cancer cases and healthy controls creates a two-gate design that can exaggerate separation relative to a symptomatic diagnostic population. Matching a few demographic variables does not reproduce the lesion spectrum of patients undergoing biopsy. Although this design is reasonable for discovery, its diagnostic performance must be interpreted within that discovery context.

Index-test concerns arise when markers and thresholds are chosen after examining the same outcome-labeled samples. A separate validation group helps only if feature selection, normalization and threshold selection remain independent of it. Unclear reporting of blinding or flow is not evidence of misconduct; it means the direction and size of potential bias cannot be resolved. The reference-standard judgments remain conservative because documented histology for cases does not resolve all questions about control classification, assessor blinding or verification timing (Table 5).

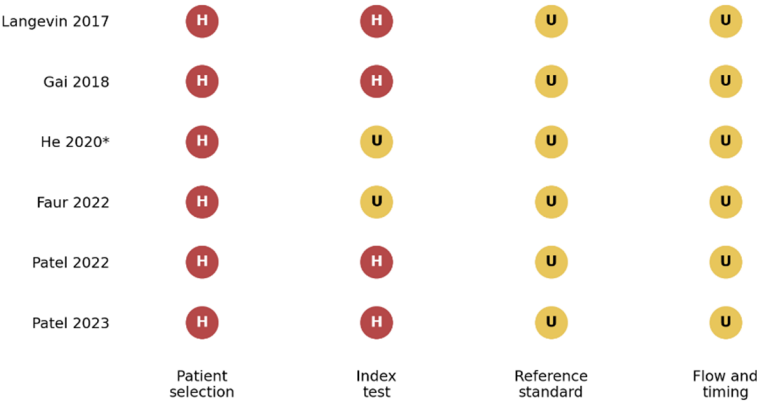


Figure 2: Provisional QUADAS-2 Assessment

H: High Risk, U: Unclear Risk

*He was assessed from the indexed abstract. All reports have substantial applicability concerns for early-OSCC screening. These are reviewer judgments, not a completed independent consensus assessment

Table 5: Provisional QUADAS 2 Judgments with Supporting Reasons

Report	Selection/index/reference/flow	Basis and applicability
Langevin [13]	High/High/Unclear/Unclear	Known cases vs controls; exploratory thresholds; mixed-site population; threshold/count narrative discrepancy.
Gai [14]	High/High/Unclear/Unclear	Two-gate recruitment and healthy controls; Youden optimization; case histology documented but reference blinding incompletely established.
He [15]	High/Unclear/Unclear/Unclear	Known preoperative cases vs normal controls; abstract insufficient for remaining domains. Assessment incomplete.
Faur [16]	High/Unclear/Unclear/Unclear	Case-control selection; mixed sites; histological diagnosis reported; 25 recruited vs 24 analyzed cancer samples needs flow clarification.
Patel 2022 [17]	High/High/Unclear/Unclear	Selected population; developmental ROC analysis; pooled controls and overlapping specimen types complicate units and flow.
Patel 2023 [18]	High/High/Unclear/Unclear	Selected cohorts; comparison of multiple panels; locked external validation not established; salivary ROC denominator unclear.

Table 6: Provisional Confidence in Clinically Relevant Claims

Claim	Confidence	Reason
A validated accuracy estimates for early OSCC	Very low	Bias, indirectness, sparse early-stage data and imprecision.
Reliable case-control discrimination by one transferable assay	Very low	Different markers and workflows; developmental accuracy and limited replication.
Panels consistently outperform single markers	Not established	No adequate independent paired clinical comparisons.
Distinguishing early OSCC from OPMDs	Insufficient direct evidence	Relevant comparator spectrum not adequately represented.
Population screening improves outcomes	No direct utility evidence identified	No retained clinical-impact screening evaluation.

A numerical quality score would obscure these different mechanisms. The traffic-light Figure 2, therefore displays domain-level provisional judgments and does not average them. He’s abstract-only assessment is visibly marked as incomplete. Laboratory reporting concerns, including vesicle characterization and normalization, are discussed alongside QUADAS-2 without pretending that every analytical limitation maps directly onto a conventional diagnostic-bias domain. Independent assessment of the full reports could alter several judgments.

Quality and Certainty of Evidence

Confidence in a precise accuracy estimate for any proposed early-OSCC screening application is very low in this provisional assessment. This conclusion reflects indirect patient selection, imprecise small cohorts, inconsistent findings and uncertainty about locked thresholds and independent validation. It does not mean that every biomarker is ineffective. Rather, the current evidence does not sufficiently constrain the expected sensitivity, specificity or clinical consequences in the intended population.

For symptomatic triage, the evidence is also insufficient to establish a safe decision rule. For population screening and for discrimination between early OSCC and clinically relevant OPMDs, direct evidence of clinical utility was not identified in the retained set. No certainty rating is assigned to a nonexistent pooled effect. Publication bias cannot be quantified with this sparse, heterogeneous evidence; selective reporting of favorable markers remains a plausible concern rather than a measured effect (Table 6-9).

Table 7: Candidates not Retained in the Diagnostic Synthesis

Report	Available source	Primary reason
Hofmann cargo study [19]	Full text	No eligible numerical miRNA diagnostic accuracy endpoint; contextual expression/cargo evidence
Hofmann paired-biofluid study [20]	Full text	No eligible numerical diagnostic endpoint; panel discovery/correlation evidence
Coon [21]	Full text	Cell-line study without patient diagnostic cohort
Jayaseelan [22]	Full text	Computational analysis without qualifying patient salivary diagnostic testing
Yap [23]	Indexed abstract	Sampling feasibility; no eligible cancer diagnostic comparison
Farag [24]	Primary bibliographic/search record and abstract information	Expression study; no verified qualifying accuracy estimate. Full-text decision provisional

Table 8 Extraction Fields for Review Completion

Category	Fields to record
Identity and linkage	Report ID; DOI; recruitment center and dates; associated publications; suspected cohort overlap.
Participants	Consecutive/random recruitment; enrolled/analyzed counts; age; sex; tobacco; oral health; stage; subsite; OPMD category.
Specimen workflow	Saliva type and volume; collection time; food/hygiene restrictions; storage; freeze-thaw cycles; contamination checks.
Analytical methods	Isolation protocol; EV characterization; miRNA identity/strand; platform; controls; normalization; detection limits.
Validation	Discovery vs validation participants; pooled samples; locked thresholds; blinding; independent site and laboratory.
Diagnostic outcomes	TP, FP, FN, TN; AUC and interval; sensitivity/specificity at fixed cutoff; early-stage subgroup counts.
Missingness and utility	Invalid tests; excluded samples; verification delays; referral changes; diagnostic delay; costs and harms.

Table 9: QUADAS 2 Completion Template

Domain	Signaling questions and evidence
Patient selection	Consecutive or random series? Case-control design avoided? Inappropriate exclusions avoided? Record population applicability.
Index test	Interpreted without reference result? Threshold prespecified? Feature selection separated from evaluation? Record assay applicability.
Reference standard	Correctly classifies target disease? Interpreted without index result? Record reference-standard applicability.
Flow and timing	Appropriate interval? Same reference standard? All included in analysis? Explain failed assays and verification differences.
Judgment record	Low, high or unclear per domain; supporting source location; initial assessor; independent assessor; resolution.

RESULTS

Discrimination and Operating Characteristics

The retained reports describe AUCs spanning 0.59 to 0.99, but this range combines different tests and populations and is not a summary measure of one technology. The lower values concern mixed-site single-marker applications, while the highest values derive from small developmental OSCC analyses. Figure 3 displays selected reported AUCs with their population labels. Only uncertainty bounds explicitly reported for the corresponding overall AUCs are shown; unavailable confidence intervals are not fabricated.

AUC does not determine the proportion of early cancers missed at a chosen threshold. Likewise, high specificity with low sensitivity may be compatible with a confirmatory research role while being unsuitable for rule-out screening. No retained dataset supports replacing biopsy or withholding investigation of a suspicious lesion after a negative salivary result. The abstract's pooled sensitivity of 0.86 and specificity of 0.85 are omitted because no reproducible underlying extraction or meta-analytic model was supplied or established here.

Early Disease and Subgroup Results

The evidence does not support a stable stage I-II sensitivity estimate. Early cases were sparse or subgroup denominators were insufficiently linked to diagnostic counts. Tumor category, overall stage, anatomical site and histological grade were sometimes discussed together but are not interchangeable. Where primary-source subgroup statements conflicted, the discrepancy was preserved rather than resolved by choosing the more favorable number.

No pooled subgroup analysis by smoking, sex, HPV status, collection method or vesicle isolation method was defensible. An apparent difference between two studies can reflect several simultaneously varying factors. Study-level comparisons cannot identify a causal explanation when sample preparation, population and analytic platform all change together. The appropriate result is unresolved heterogeneity, accompanied by specific variables for future extraction and validation.

Panels and Clinical Utility

The three-miRNA panel showed striking developmental discrimination, but the evidence does not establish that panels consistently outperform single markers across independent clinical cohorts. Demonstrating superiority requires a paired comparison, the same reference standard and clinically relevant operating points. A higher apparent AUC from a more flexible model may reflect true complementary information, overfitting or both. None of these possibilities can be distinguished by comparing the largest AUCs across unrelated studies.

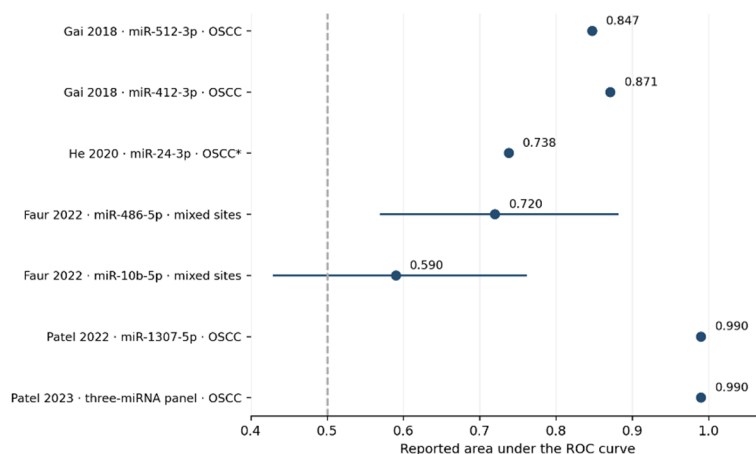


Figure 3: Selected Study-Reported AUCs.

Points are descriptive and are not pooled or directly comparable. Lines show reported 95% confidence intervals only for Faur's overall estimates. Other intervals were unavailable in the extracted material. *He: abstract-only extraction. Patel estimates require denominator clarification. No synthetic ROC curves were created.

No retained report demonstrated reduced diagnostic delay, fewer missed cancers, improved survival, acceptable downstream biopsy burden or cost-effectiveness from a salivary-miRNA testing strategy. These outcomes are distinct from laboratory discrimination. The absence of such results limits practical recommendations even where analytical findings are encouraging. It also identifies the next evidentiary step: prospective evaluation of a specified assay embedded in a specified clinical pathway.

DISCUSSION

Principal Interpretation

The central finding is a mismatch between biological promise and the level of clinical validation. Several candidate microRNAs separate selected cancer samples from controls, but results are not consistent enough to define a standardized early-detection test. Very high AUCs are concentrated in small developmental settings where denominator interpretation and model independence require clarification. More moderate or weak discrimination appears in other reports, including studies of frequently cited markers. A defensible conclusion must accommodate both patterns.

The distinction between detection and prognosis further narrows interpretation. A marker associated with aggressive disease can be useful for prognosis while being insensitive to small lesions. Conversely, a marker measurable in early disease may not discriminate cancer from non-malignant mucosal changes. These are different tasks requiring different sampling designs. The term "early detection" should be reserved for a defined stage or clinical time point and supported by results specific to that target.

Comparison With Earlier Reviews

Faur's 2021 review retained five studies across a broader head-and-neck and premalignant-lesion scope, whereas Sanesi's 2024 review included nine articles and described 41 differentially expressed microRNAs [26,27]. Those counts are not directly comparable with six diagnostic reports in the present restricted assessment because eligibility, search coverage and outcome definitions differ. In particular, a report of altered expression or an OLP biomarker may enter a broad review while failing a diagnostic-accuracy criterion for OSCC.

The current manuscript contributes an explicit separation between diagnostic estimates, expression findings and clinical utility. It also records discrepancies rather than converting favorable narrative conclusions into extracted data. Nonetheless, it cannot claim superior comprehensiveness: the earlier reviews used broader source coverage and this workflow has known access limitations. A final dissertation review should use their reference lists for additional retrieval while independently reassessing eligibility and avoiding duplication of their conclusions without checking primary sources.

Biological Plausibility and Tissue Attribution

Concordance between tissue expression, salivary vesicle cargo and experimental effects can make a biomarker hypothesis more credible. It does not show that the measured salivary signal arises exclusively from malignant cells. Whole-mouth saliva samples a complex biological environment. Even a signal related to cancer may reflect host response, inflammation or tumor burden rather than a transformation-specific molecular event. These alternatives lead to different expectations for specificity and stage dependence.

Mechanistic experiments answer whether manipulating a microRNA can affect cellular behavior under the tested conditions. Diagnostic studies answer whether measuring it correctly classifies patients. Neither question can replace the other. A biologically active microRNA with broad expression might be a poor discriminator, while a useful classifier need not be a causal driver. Keeping these roles separate prevents pathway enrichment, cell proliferation assays or genomic survival analyses from being treated as independent diagnostic validation.

Preanalytical and Laboratory Heterogeneity

The assay includes the collection and isolation workflow, not merely the miRNA name. Saliva stimulation, food restrictions, oral hygiene, time of day, storage, freeze–thaw exposure and removal of cellular material can alter the material entering analysis. Clinically deployable test needs evidence that its decision rule remains stable across permissible variation. Standardizing procedures only after marker selection may require re-estimation of thresholds because the measurement scale can change.

Boulestreau and colleagues directly examined how salivary vesicle isolation methods affect downstream biomarker detection [28]. Their work supports treating isolation method as part of assay identity. Precipitation, centrifugation and other enrichment methods recover overlapping but non-identical preparations. An analytically convenient method may be useful, but purity, recovery and reproducibility need to be quantified rather than inferred from the kit label. MISEV-compatible characterization should accompany diagnostic development.

Normalization introduces a separate source of variability. An endogenous reference must remain stable in the intended disease and control groups; an exogenous spike-in controls only selected processing steps. The MIQE guidelines provide a framework for reporting qPCR procedures, controls and analytical performance [29]. Future comparisons should report failed reactions, detection limits and indeterminate results. Excluding difficult samples without reporting their frequency can make the apparent accuracy exceed the performance of the complete testing process.

Statistical Interpretation and Validation

Small datasets with many candidate features are especially vulnerable to optimistic model performance. Candidate selection, transformation, normalization, threshold selection and panel construction can all use outcome information. Separating a “validation” cohort is insufficient if repeated model choices depend on its results. A defensible validation study should lock the assay and decision rule beforehand, record all eligible patients and report performance without additional optimization.

Near-perfect percentages require examination of the underlying denominators. With five independent controls, observed empirical specificity changes in increments of twenty percentage points. A reported specificity of 99% therefore cannot be interpreted as a simple control classification fraction from five people. It may originate from fitted curves or a different analytical unit, but that explanation must be documented. This arithmetic observation does not invalidate the entire experiment; it limits which quantities can responsibly enter a diagnostic meta-analysis.

Pooling cannot repair these problems. A random-effects model allows genuine variation across compatible studies; it does not make different targets, specimens or outcomes interchangeable. Combining miRNAs because they share an exosomal carrier would resemble combining unrelated laboratory tests into one sensitivity estimate. Marker-specific meta-analysis may eventually be appropriate, but only after test definitions, thresholds, populations and participant independence are sufficiently clear.

Implications for Practice and Policy

The present evidence supports research use and structured validation. It does not justify using a salivary-miRNA result to defer assessment of a suspicious lesion. The first practical question should be whether the test adds useful information to examination and existing referral criteria. A costly molecular assay that reproduces an already obvious clinical distinction may have limited value, even when its AUC is favorable.

For health systems, inexpensive specimen collection should not be equated with inexpensive testing. Vesicle processing, RNA extraction, equipment maintenance, quality assurance and repeat testing contribute to the complete pathway. Policy evaluation should include laboratory failures, turnaround time, access and downstream investigations. These requirements are particularly relevant where referral and biopsy capacity are constrained. A test should be judged by the consequences of the decisions it changes, rather than by analytical novelty alone.

Strengths and Limitations

A strength of this work is traceability: quantitative statements are linked to identifiable reports, the actual search denominator is recorded and unverified pooled claims are removed. The synthesis separates oral-cavity evidence from mixed-site evidence, avoids adding tissue and saliva sample counts and highlights numerical inconsistencies. Original figures summarize the evidence without redrawing hypothetical ROC curves or presenting fabricated patient data.

The principal limitations are restricted database coverage, retrospective methods, single-workflow screening and incomplete full-text access. The publication window ends in March 2025 and therefore cannot support a claim to cover all evidence available in September 2026. Selection could have missed eligible reports that use different terminology or appear outside Europe PMC. Several assessments remain provisional and publication-level results cannot resolve patient overlap or missing diagnostic counts. These limitations require completion, not merely acknowledgment, before submission as a comprehensive systematic review.

Clinical Significance and Practical Implications

Salivary collection may reduce sampling burden and enable repeat measurements, but its practical value depends on a validated clinical use. The most defensible initial application to investigate is adjunctive triage in a clearly defined lesion population. Such studies should record what clinicians would do without the assay and whether the assay changes that decision appropriately. An observed molecular change alone does not establish a benefit for patients.

Clinical teams should interpret exploratory microRNA findings as research evidence. Suspicious lesions still require appropriate clinical investigation regardless of an experimental assay result. Future reports should translate accuracy into expected false-positive and false-negative consequences at the prevalence of the intended setting, using externally validated inputs. This review does not provide those forecasts because the necessary target-setting accuracy estimates have not been established.

Future Research Recommendations

The next priority is a prospective, multicenter diagnostic cohort recruited consecutively before definitive diagnosis. It should include early OSCC, clinically relevant OPMDs, benign lesions and inflammatory controls, with prespecified anatomical boundaries. The reference-standard process, timing and handling of patients without biopsy should be documented. Eligibility should not depend on whether a sample is easy to analyze. Separate reporting of stage I-II performance is essential if early detection is the intended claim.

The analytical workflow and classifier should be locked before external testing. Laboratories should report vesicle enrichment, characterization, RNA yield, normalization, reproducibility and prespecified quality-control failure rules. Patient-level predictions and two-by-two tables should be available at each clinically relevant threshold. Repeated measurements, pooled controls and technical replicates must be identified so that the patient remains the unit of diagnostic inference. Sample size should be chosen around precision for the critical sensitivity and specificity targets.

Finally, direct comparisons of the best single marker, a prespecified panel and clinical examination should be performed in the same participants. Only after successful external validation should investigators evaluate implementation outcomes, biopsy burden and cost-effectiveness. Separate studies are needed for prognosis, surveillance and OPMD progression because those questions require longitudinal outcomes. For this manuscript, completion priorities are the unexecuted database searches, independent screening, full-text retrieval and author clarification of disputed diagnostic denominators.

CONCLUSIONS

Salivary extracellular vesicle-associated microRNAs provide plausible candidates for OSCC detection, but the verified literature remains developmental. Reported discrimination varies substantially and the highest estimates require external validation and clearer patient-level reporting. Evidence specifically supporting early-stage detection or clinical screening utility is insufficient. No pooled sensitivity, specificity or universally superior panel is justified by this assessment. The appropriate next step is rigorous validation of defined assays in clinically representative populations, accompanied by completion of the systematic review methods documented here.

Declarations

Registration: No prospective registration is claimed. Ethics and consent: This manuscript analyzes published reports and contains no newly collected patient data. Funding, competing interests, author contributions and corresponding-author details: To be supplied and confirmed by the authors. Data availability: Source-level extraction, search strategies and assessment templates are included below. AI assistance: Artificial intelligence assisted literature retrieval, drafting, formatting and consistency checks; author verification of all evidence, eligibility decisions and final text is required. No simulated patient data or fabricated meta-analysis was used.

REFERENCES

- [1] Bray, F. 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*, vol. 74, no. 3, 2024, pp. 229–263. <https://doi.org/10.3322/caac.21834>.
- [2] Lydiatt, W.M. et al. "Head and neck cancers—Major changes in the American joint committee on cancer eighth edition cancer staging manual." *CA: A Cancer Journal for Clinicians*, vol. 67, no. 2, 2017, pp. 122–137. <https://doi.org/10.3322/caac.21389>.

- [3] Park, N.J. et al. "Salivary microRNA: Discovery, characterization and clinical utility for oral cancer detection." *Clinical Cancer Research*, vol. 15, no. 17, 2009, pp. 5473–5477. <https://doi.org/10.1158/1078-0432.CCR-09-0736>.
- [4] Théry, C. et al. "Minimal information for studies of extracellular vesicles 2018 (MISEV2018): A position statement of the international society for extracellular vesicles and update of the MISEV2014 Guidelines." *Journal of Extracellular Vesicles*, vol. 7, no. 1, 2018, pp. 1535750. <https://doi.org/10.1080/20013078.2018.1535750>.
- [5] Welsh, J.A. et al. "Minimal information for studies of extracellular vesicles (MISEV2023): From basic to advanced approaches." *Journal of Extracellular Vesicles*, vol. 13, no. 2, 2024, pp. e12404. <https://doi.org/10.1002/jev2.12404>.
- [6] Page, M.J. et al. "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews." *BMJ*, vol. 372, 2021, pp. n71. <https://doi.org/10.1136/bmj.n71>.
- [7] McInnes, M.D.F. et al. "Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement." *JAMA*, vol. 319, no. 4, 2018, pp. 388–396. <https://doi.org/10.1001/jama.2017.19163>.
- [8] Rethlefsen, M.L. et al. "PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews." *Systematic Reviews*, vol. 10, no. 1, 2021, pp. 39. <https://doi.org/10.1186/s13643-020-01542-z>.
- [9] Whiting, P.F. et al. "QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies." *Annals of Internal Medicine*, vol. 155, no. 8, 2011, pp. 529–536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>.
- [10] Bossuyt, P.M. et al. "STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies." *BMJ*, vol. 351, 2015, pp. h5527. <https://doi.org/10.1136/bmj.h5527>.
- [11] Schünemann, H.J. et al. "Grading quality of evidence and strength of recommendations for diagnostic tests and strategies." *BMJ*, vol. 336, no. 7653, 2008, pp. 1106–1110. <https://doi.org/10.1136/bmj.39500.677199.ae>.
- [12] Reitsma, J.B. et al. "Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews." *Journal of Clinical Epidemiology*, vol. 58, no. 10, 2005, pp. 982–990. <https://doi.org/10.1016/j.jclinepi.2005.02.022>.
- [13] Langevin, S. et al. "Comprehensive microRNA-sequencing of exosomes derived from head and neck carcinoma cells in vitro reveals common secretion profiles and potential utility as salivary biomarkers." *Oncotarget*, vol. 8, no. 47, 2017, pp. 82459–82474. <https://doi.org/10.18632/oncotarget.19614>.
- [14] Gai, C. et al. "Salivary extracellular vesicle-associated MIRNAS as potential biomarkers in oral squamous cell carcinoma." *BMC Cancer*, vol. 18, no. 1, 2018, pp. 439. <https://doi.org/10.1186/s12885-018-4364-z>.
- [15] He, L. et al. "Salivary Exosomal miR-24-3p serves as a potential detective biomarker for oral squamous cell carcinoma screening." *Biomedicine & Pharmacotherapy*, vol. 121, 2020, pp. 109553. <https://doi.org/10.1016/j.biopha.2019.109553>.
- [16] Faur, C.I. et al. "Salivary exosomal microRNA-486-5p and MicroRNA-10b-5p in oral and oropharyngeal squamous cell carcinoma." *Medicina (Kaunas)*, vol. 58, no. 10, 2022, pp. 1478. <https://doi.org/10.3390/medicina58101478>.
- [17] Patel, A. et al. "Salivary Exosomal miRNA-1307-5p Predicts disease aggressiveness and poor prognosis in oral squamous cell carcinoma patients." *International Journal of Molecular Sciences*, vol. 23, no. 18, 2022, pp. 10639. <https://doi.org/10.3390/ijms231810639>.
- [18] Patel, A. et al. "A Novel 3-miRNA network regulates tumour progression in oral squamous cell carcinoma." *Biomarker Research*, vol. 11, no. 1, 2023, pp. 64. <https://doi.org/10.1186/s40364-023-00505-5>.
- [19] Hofmann, L. et al. "Cargo and functional profile of saliva-derived exosomes reveal biomarkers specific for head and neck cancer." *Frontiers in Medicine*, vol. 9, 2022, pp. 904295. <https://doi.org/10.3389/fmed.2022.904295>.
- [20] Hofmann, L. et al. "Comparison of plasma- and saliva-derived exosomal mirna profiles reveals diagnostic potential in head and neck cancer." *Frontiers in Cell and Developmental Biology*, vol. 10, 2022, pp. 971596. <https://doi.org/10.3389/fcell.2022.971596>.
- [21] Coon, J., K. Kingsley and K.M. Howard. "miR-365 (microRNA): Potential biomarker in oral squamous cell carcinoma exosomes and extracellular vesicles." *International Journal of Molecular Sciences*, vol. 21, no. 15, 2020, pp. E5317. <https://doi.org/10.3390/ijms21155317>.
- [22] Jayaseelan, V.P. and P. Arumugam. "Exosomal microRNAs targeting TP53 gene as promising prognostic markers for head and neck squamous cell carcinoma." *Global Medical Genetics*, vol. 9, no. 4, 2022, pp. 277–286. <https://doi.org/10.1055/s-0042-1758204>.
- [23] Yap, T. et al. "Oral swirl samples—A robust source of microRNA protected by extracellular vesicles." *Oral Diseases*, vol. 23, no. 3, 2017, pp. 312–317. <https://doi.org/10.1111/odi.12603>.
- [24] Farag, A.F. et al. "MicroRNA-134/MicroRNA-200a derived salivary exosomes are novel diagnostic biomarkers of oral squamous cell carcinoma." *Egyptian Dental Journal*, vol. 67, no. 1, 2021, pp. 367–377. <https://doi.org/10.21608/edj.2020.47990.1317>.
- [25] Byun, J.S. et al. "Diagnostic profiling of salivary exosomal microRNAs in oral lichen planus patients." *Oral Diseases*, vol. 21, no. 8, 2015, pp. 987–993. <https://doi.org/10.1111/odi.12374>.
- [26] Faur, C.I. et al. "Salivary exosomal microRNAs as biomarkers for head and neck cancer detection—A literature review." *Maxillofacial Plastic and Reconstructive Surgery*, vol. 43, no. 1, 2021, pp. 19. <https://doi.org/10.1186/s40902-021-00303-9>.
- [27] Sanesi, L. et al. "Salivary exosomal microRNA profile as biomonitoring tool for diagnosis and prognosis of patients with head and neck squamous cell carcinoma: A systematic review." *Archives of Oral Biology*, vol. 165, 2024, pp. 106012. <https://doi.org/10.1016/j.archoralbio.2024.106012>.
- [28] Boulestreau, J. et al. "Salivary extracellular vesicles isolation methods impact the robustness of downstream biomarkers detection." *Scientific Reports*, vol. 14, no. 1, 2024, pp. 31233. <https://doi.org/10.1038/s41598-024-82488-3>.
- [29] Bustin, S.A. et al. "The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments." *Clinical Chemistry*, vol. 55, no. 4, 2009, pp. 611–622. <https://doi.org/10.1373/clinchem.2008.112797>.