

Review

Not peer-reviewed version

SalivaSense 2.0: Multiplex cfDNA-cfRNA Epigenome Signatures Enabling Pico-Stage Oral Squamous Cell Carcinoma Detection, Validation Frameworks and Translational Pipelines

[Abdullah Ayad](#)^{*} and Yousef Husam

Posted Date: 15 September 2025

doi: 10.20944/preprints202509.1251.v1

Keywords: Salivary biopsy; cfDNA methylation; cfRNA profiling; Microfluidic extraction; OSCC detection



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

SalivaSense 2.0: Multiplex cfDNA-cfRNA Epigenome Signatures Enabling Pico-Stage Oral Squamous Cell Carcinoma Detection, Validation Frameworks and Translational Pipelines

Abdullah Ayad ^{1,2,*} and Yousef Husam ¹

¹ College of Dentistry, University of Basrah, Basra, 61004, Iraq

² Molecular Genetics and Proteomics Department, Bayan National Lab for Advanced Medical Diagnostics, Basra, 61004, Iraq

* Correspondence: abdullah.ayad@bayangen.com or dent23200021@avicenna.uobasrah.edu.iq; Tel: +9647853494170

Abstract

Background: Oral squamous cell carcinoma (OSCC) is characterized by a high mortality rate primarily because it is often diagnosed at an advanced stage, creating a profound global survival disparity. Early detection is paramount, yet a significant screening gap persists, especially in community and low-resource dental settings. This review articulates the case for SalivaSense 2.0, a conceptual framework designed to transform routine saliva sampling into an ultrasensitive multiplex liquid biopsy capable of detecting the disease at its nascent, "pico-stage". **Main Body:** The core innovation of SalivaSense 2.0 lies in the synergistic integration of stable cell-free DNA methylation signatures with dynamic cell-free RNA transcriptional profiles. This dual-analyte approach generates a comprehensive epigenomic fingerprint that can characterize tumorigenesis at its earliest detectable stages. The proposed assay is built on an automated microfluidic cartridge for highly efficient nucleic acid isolation from saliva. Downstream analysis employs advanced technologies like nanopore sequencing and electrochemical nanobiosensors. This system is unified by a sophisticated artificial intelligence pipeline that uses molecular barcoding and neural networks to suppress sequencing errors and computationally mitigate biological noise from confounders such as tobacco use, periodontal inflammation, and the oral microbiome. This achieves an exceptionally low detection limit, enabling the identification of minute cancer signals. The framework also delineates a complete translational pathway, from rigorous analytical and clinical validation benchmarks to the development of accessible point-of-care devices, including smartphone-linked readers and simplified sample logistics. **Conclusion:** While significant challenges in clinical validation, global data harmonization, and demonstrating economic feasibility remain, overcoming these obstacles is essential. Successfully translating this technology from the laboratory to the front lines of care could decisively shift OSCC management from reactive surgical intervention toward proactive, saliva-based preventive screening, dramatically improving patient outcomes worldwide.

Keywords: Salivary biopsy; cfDNA methylation; cfRNA profiling; Microfluidic extraction; OSCC detection

1. Background

Oral squamous cell carcinoma (OSCC) represents a profound and escalating global health challenge [1–16]. As the most prevalent malignancy of the head and neck, it accounts for more than 90% of all oral cancers [1,6,8] and is responsible for an estimated 389,000 new cases and 188,000 deaths annually worldwide [6]. This burden is not evenly distributed; from 1990–2021, the absolute number

of deaths and disability-adjusted life years (DALYs) from lip and oral cavity cancer increased by 92.92% and 113.94%, respectively (Figure 1) [5]. This growing crisis is marked by stark sociodemographic disparities, with mortality and DALY rates consistently highest in low- and middle-sociodemographic index (SDI) regions, even when incidence rates are higher in more developed countries [16]. These statistics paint a sobering picture of a disease where geography and economic status are potent determinants of survival [4,16].

The lethality of OSCC is tragically linked to a marked drop in survival when the disease is not caught in infancy [3,8]. The difference between early- and late-stage diagnosis is a matter of life and death: the five-year survival rate can reach 70–90% for localized, early-stage tumors but decreases to less than 50% once the cancer has progressed [6,8]. In some populations, the five-year survival for advanced disease is a staggering 20%, compared with 80% for stage I cancers [10]. The core of the problem lies in delayed detection. The vast majority of patients, over 85% in some studies, are diagnosed at an advanced stage (III or IV) [10]. This delay is particularly pronounced in low- and middle-income countries, where 60–70% of patients present with late-stage disease, compared with 30–40% in high-resource settings [6]. This diagnostic delay not only decreases survival rates but also has devastating consequences for quality of life, with extensive surgeries leading to permanent functional impairments in speech and swallowing [6].

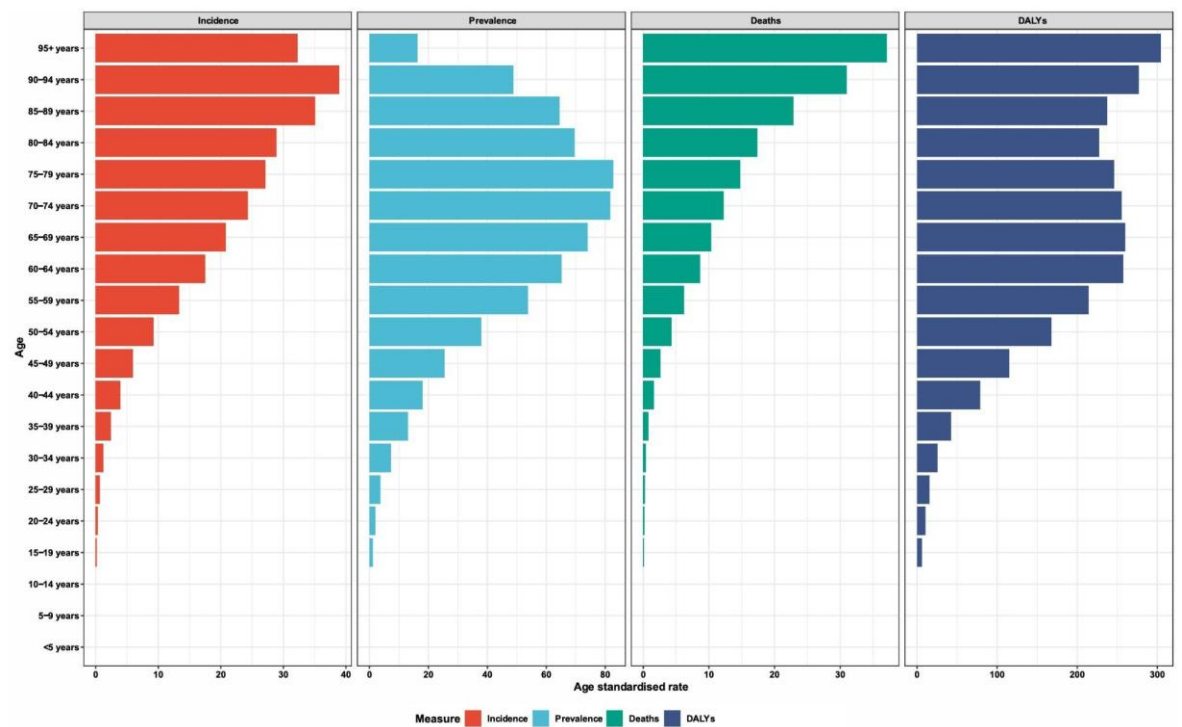


Figure 1. The global epidemiological burden of oral cancer by Age-standardized rates. Age-standardized rates for four key public health measures: incidence (red), prevalence (light blue), deaths (green), and disability-adjusted life years (DALYs, dark blue). The data illustrate a dramatic increase in all metrics with advancing age, highlighting the significant burden of the disease on aging populations. Reprinted from [5] under the terms of the Creative Commons Attribution 4.0 International License.

This diagnostic latency stems from a critical screening void, especially within community and low-resource dental settings, which should serve as the first line of defense [2,4]. Dental professionals conduct more than 75% of all oral cancer screenings, making routine dental visits an unparalleled opportunity for early detection [4]. However, this opportunity is frequently missed. Access to dental care is a significant barrier, particularly for low-income adults, who bear the greatest burden of the disease [4]. A landmark study demonstrated that the elimination of Medicaid dental benefits in California was associated with a 6.5%-point decline in early-stage diagnoses of oral cavity cancers,

powerfully illustrating the link between dental access and survival [4]. Even when patients do see a dentist, challenges persist. Many practitioners in developing countries feel that they lack the time, training, and confidence to conduct screenings effectively, with fewer than half feeling adequately knowledgeable [2,11]. This gap creates a perfect storm where the disease goes unnoticed until it is far too late, resulting in not only loss of life but also a substantial economic burden. The cost of treating advanced-stage OSCC can be 22% to 373% higher than that of treating early-stage disease, representing a catastrophic economic burden on healthcare systems and individuals alike [14].

1.1. Saliva as a Precision Liquid Biopsy

In the quest to bridge this deadly diagnostic gap, the concept of "liquid biopsy", the analysis of tumor-derived materials in bodily fluids, has emerged as a transformative alternative to invasive tissue biopsies [7,8]. While blood has been the traditional focus, saliva offers a uniquely powerful, noninvasive window into the molecular landscape of OSCC [7,8,15]. The anatomical proximity of the oral cavity to the salivary glands allows for direct shedding of tumor components, including circulating tumor nucleic acids (ctNAs), extracellular vesicles (EVs), and exfoliated tumor cells [7,8]. This localized shedding may lead to a higher concentration and more sensitive detection of biomarkers compared with those in blood [9,15]. The collection process itself is a patient-centered approach: it is painless, inexpensive, can be performed without specialized personnel, and allows for the painless repeated sampling necessary for real-time monitoring of disease progression and treatment response [7,8]. The overall workflow for salivary liquid biopsy involves the collection of saliva, followed by the detection and analysis of various tumor-derived biomarkers, including ctDNA, ctRNA, and extracellular vesicles, which then inform a range of clinical applications from early diagnosis to therapeutic monitoring (Figure 2) [7].

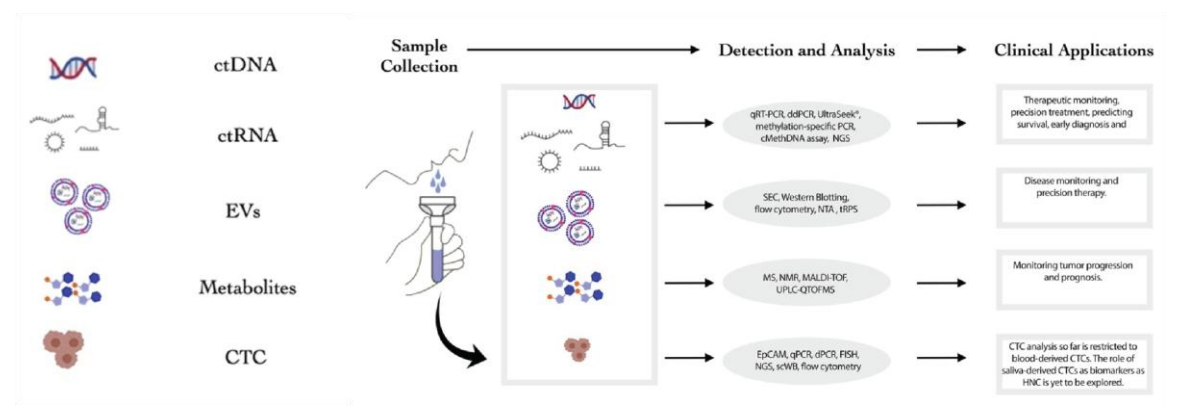


Figure 2. Conceptual workflow of salivary liquid biopsy for clinical applications. The process begins with noninvasive sample collection. Key analytes, including circulating tumor DNA (ctDNA), circulating tumor RNA (ctRNA), extracellular vesicles (EVs), metabolites, and circulating tumor cells (CTCs), are then isolated. Specialized detection methods are applied to each analyte, yielding data that can be used for therapeutic monitoring, disease progression analysis, and early diagnosis of oral cancers. Reprinted from [7] under the terms of the Creative Commons Attribution 4.0 International License.

The true power of salivary liquid biopsy lies in its ability to provide a multilayered, synergistic view of the tumor by analyzing both cell-free DNA (cfDNA) and cell-free RNA (cfRNA). While cfDNA reveals the stable genetic and epigenetic blueprint of the tumor, including cancer-specific mutations and DNA methylation patterns, cfrRNA offers a dynamic, real-time snapshot of the tumor's functional state through its gene expression profile [7–9,15]. One study demonstrated that for cfrRNA analysis, saliva is a more effective medium for OSCC detection than blood plasma is [15]. This synergy is crucial; combining the stable, cancer-specific "what" from cfDNA with the dynamic "how"

from cfRNAs provides a far more robust and comprehensive signature of malignancy than either molecule alone could alone.

The evolution of salivary biomarkers reflects a continuous journey toward greater precision and clinical utility. Initial research focused on proteins, enzymes, and metabolites, which, while valuable, often lacked the specificity needed for definitive diagnosis owing to their presence in various physiological and noncancerous pathological states [1,3,7]. The field then advanced into the "-omics" era, first exploring the transcriptomes of dysregulated microRNAs (miRNAs) and messenger RNAs (mRNAs) [7,8]. More recently, the focus has shifted to the epigenome, particularly the analysis of DNA methylation patterns [3,9]. These epigenetic marks are highly stable and often tumor specific, offering a robust signal that can be detected even in the earliest stages of carcinogenesis [9]. It is at this cutting edge that a technology such as SalivaSense 2.0 is positioned. It represents the logical next step in this evolution, moving beyond single-analyte discovery to a multiplexed approach that integrates the most stable and specific signals from both the cfDNA epigenome and the cfRNA transcriptome. This integrated strategy aims to create a highly sensitive and specific signature capable of detecting OSCC at its inception, transforming early diagnosis from a challenge into a clinical reality.

The grim reality of OSCC is that it remains a devastatingly lethal disease, not because it is untreatable but because it is often found too late [6,8,10]. A vast and inequitable screening void persists precisely in low-resource and community settings, where early detection could have the greatest impact [4,11,12,16]. While salivary diagnostics have shown immense promise, the translation from laboratory discovery to clinical application has been fraught with challenges related to sensitivity, specificity, and standardization [3,7,8]. The next frontier in this field must therefore focus on creating a robust, validated, and accessible diagnostic framework that can overcome these historical barriers.

This review synthesizes the current evidence to build the case for SalivaSense 2.0, a hypothetical next-generation diagnostic framework predicated on multiplexed salivary cfDNA and cfRNA epigenome signatures. We aim to delineate a comprehensive roadmap for its development, validation, and implementation for the pico-stage detection of OSCC, catching the disease at its molecular dawn. To achieve this goal, this paper will be guided by three central questions. First, what is the multifaceted global burden of OSCC, and why does the chasm between early- and late-stage survival mandate a paradigm shift toward pico-stage detection? Second, how has saliva evolved into a precision liquid biopsy medium, and what is the synergistic potential of combining cfDNA and cfRNA epigenomic signatures to overcome the limitations of previous single-analyte biomarkers? Finally, what are the critical components of a translational pipeline, from robust validation frameworks to equitable implementation strategies, required to move a technology such as SalivaSense 2.0 from the research bench to the front lines of community health and global oncology? Answering these questions requires an integrative understanding of OSCC biomarkers and the convergent technologies used for their detection and analysis. The subsequent sections of this review delve into key technological domains, including advanced sequencing, biosensors, and artificial intelligence, that form the foundation of next-generation salivary diagnostics (Figure 3) [7].

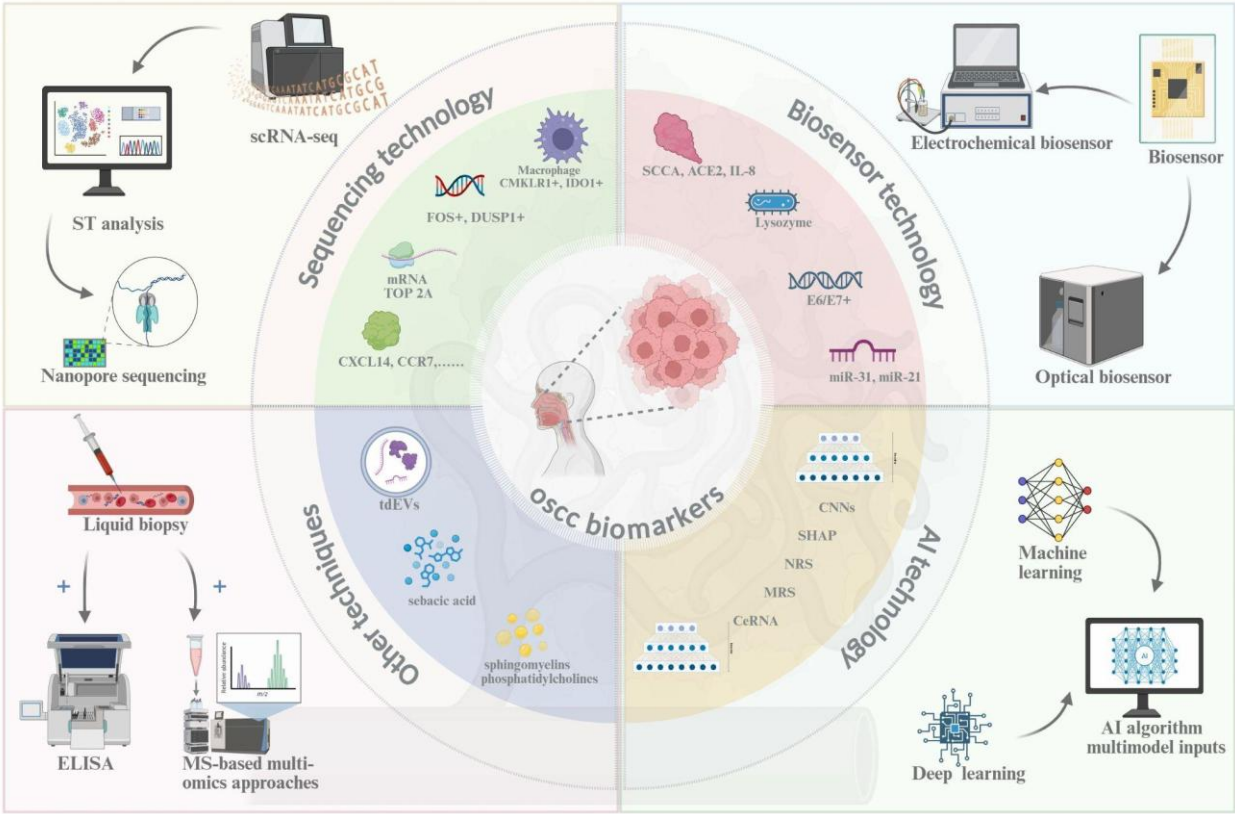


Figure 3. An Integrated View of Modern Technologies and Biomarkers for OSCC Detection. The ecosystem for OSCC diagnostics is composed of four key technological pillars: sequencing technology (e.g., scRNA-seq, nanopore), biosensor technology (e.g., electrochemical and optical sensors), AI technology (e.g., machine learning and deep learning), and other techniques (e.g., liquid biopsy, ELISA). These technologies are used to detect a wide array of salivary OSCC biomarkers, including nucleic acids (mRNA, miRNA), proteins (SCCA, lysozyme), and metabolites, to enable early and precise diagnosis. Reprinted from [6] under the terms of the Creative Commons Attribution 4.0 International License.

2. Molecular and Biophysical Findings

The advent of salivary liquid biopsy has led to a deep understanding of the molecular processes that governs the life of cell-free nucleic acids (cfNAs) [17–30]. These fleeting molecular fragments, which are ejected from cells throughout the body, embark on a remarkable journey, carrying with them a dense repository of biological information that mirrors the genomic and epigenomic state of their parent tissues [19,22]. To harness their diagnostic power, particularly for detecting nascent oral squamous cell carcinomas (OSCCs), we must first unravel the fundamental principles of their creation, their transport into the oral cavity, and the stabilizing associations that shields them from degradation. This exploration reveals a world where the mechanisms of cell death, the architecture of our own chromatin, and the enzymatic landscape of saliva converge to shape the very nature of these powerful biomarkers.

2.1. Biogenesis, Release Kinetics and Salivary Trafficking

The story of a cfDNA or cfRNA molecule begins at its moment of release, a process governed by the life and death of a cell [22]. In the context of oral carcinogenesis, the dysplastic epithelium becomes a primary source, shedding these nucleic acid fragments into the surrounding biofluid through a combination of passive leakage and active, vesicular export [18,20]. The dominant release pathways are processes of cell death, including apoptosis and necrosis [19,30]. Apoptosis, a form of programmed cell death, involves the orderly cleavage of DNA by endonucleases into fragments that reflect the fundamental packaging of our genome [19,22]. This process generates the characteristic

mononucleosomal cfDNA peak at approximately 167 base pairs (bp), a length corresponding to the DNA wrapped around a core of histone proteins plus a small linker segment [20,22,25]. In contrast, necrosis, a more chaotic form of cell death often observed in tumors, results in a more haphazard release of DNA, yielding longer and more variably sized fragments that can extend up to 10 kilobases (kb) or more [20,22].

In addition to being passively released, cells also actively secrete cfNAs, which often package them within extracellular vesicles (EVs), such as exosomes and microvesicles [18,28]. These lipid-bound nanocarriers are produced during normal cell metabolism and are crucial for intercellular communication, ferrying their cargo of DNA, RNA, and proteins between cells [18,28]. In cancer, this process is coopted, with tumor cells releasing EVs that carry tumor-specific genetic and epigenetic alterations [20,30]. This dual-origin narrative, which involves passive shedding from dying cells and active export from living cells, creates a complex and information-rich pool of cfNAs.

Once released, these nucleic acids must traverse the oral cavity to become salivary biomarkers. Saliva is a unique biofluid in that it captures cfNAs from two distinct sources: local and systemic [17,20]. Locally, cfDNA and cfRNA are shed directly from the tissues of the oral cavity, including the tumor itself and surrounding epithelial cells [20,30]. Systemically, cfDNA from distant sites, such as hematopoietic cells, which are the primary source in healthy individuals, travels through the bloodstream and enters the saliva via passive diffusion and is transported through capillary-rich salivary glands [17,23,30]. This dual trafficking mechanism makes saliva a powerful diagnostic medium, offering a direct sample of the tumor microenvironment while also providing a window into systemic disease processes [20].

The journey from the cell to the saliva imprints a distinct signature onto the cfDNA fragments themselves. The resulting population of salivary cfDNA (scfDNA) exhibits a characteristic fragmentation profile that differs from that found in blood plasma [17]. Analysis of scfDNA revealed a peculiar, "jagged" size distribution with multiple peaks between 35 and 300 bp, in stark contrast to the more defined mononucleosomal peak observed in plasma [17]. This jagged pattern is a direct reflection of how DNA is packaged within the cell nucleus. The surviving fragments are essentially "nucleosome footprints," which are remnants of DNA that are protected from enzymatic degradation by being tightly wound around histone proteins or bound by transcription factors [19,25]. The locations and sizes of these fragments thus provide a map of the chromatin landscape of the cell of origin, revealing which genes were active or silent at the time of release [19,25].

This principle is clearly visualized in the characteristic fragmentation profiles observed in the cfDNA analysis, which show distinct peaks corresponding to the fundamental packaging units of the genome (Figure 4) [19].

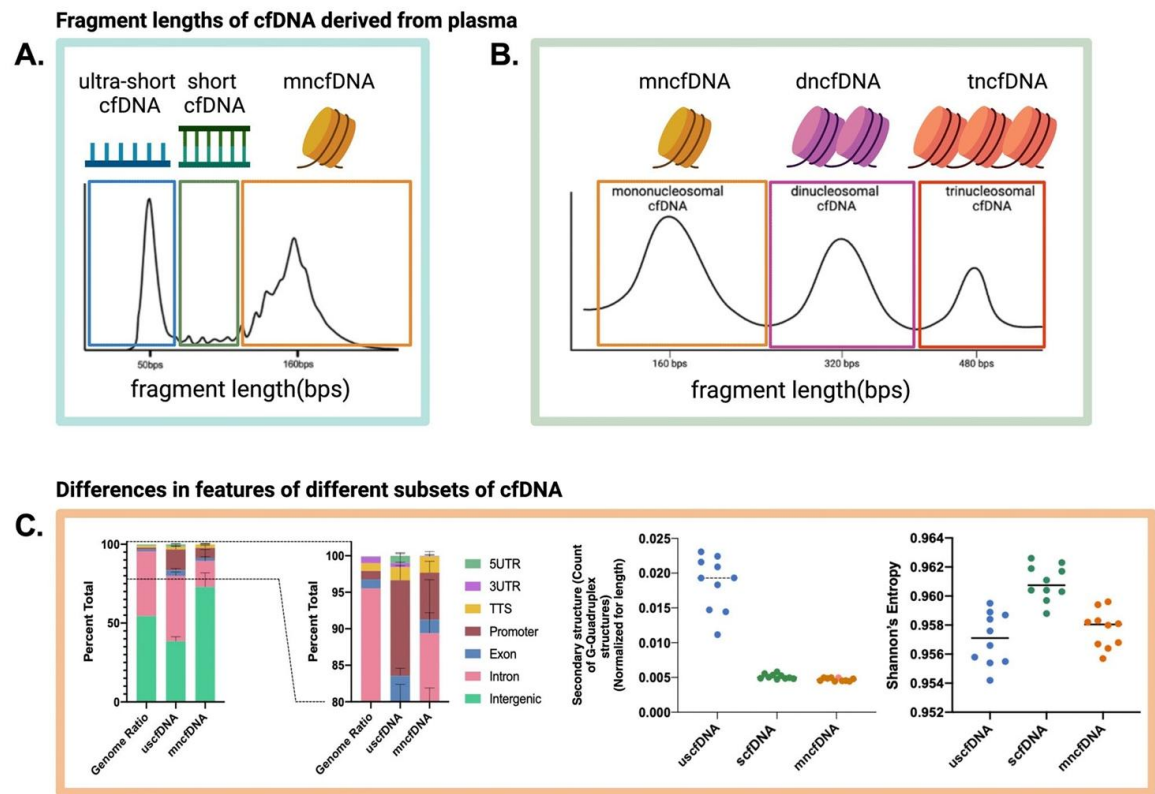


Figure 4. Characteristic fragmentation profiles of cell-free DNA (cfDNA). (A) A representative size distribution profile of cell-free DNA (cfDNA) fragments isolated from plasma. The profile shows two prominent populations: ultrashort cfDNA (uscfDNA), which peaks below 100 bp (typically ~50 bp), and mononucleosomal cfDNA (mnucfDNA), which has a characteristic peak at approximately 167 bp. (B) Schematic representation and idealized fragment length distributions for cfDNA originating from mononucleosomes (mnucfDNA, ~160 bp), dinucleosomes (dnucfDNA, ~320 bp), and trinucleosomes (tnucfDNA, ~480 bp), illustrating the ladder pattern of nucleosome-protected DNA. (C) Comparative analysis of cfDNA features. Left: The genomic origins of uscfDNA and mnucfDNA differ, with uscfDNA showing higher proportions of fragments from promoter and exon regions than mnucfDNA and the expected genome ratio. Middle: uscfDNA displays a significantly higher normalized count of G-quadruplex secondary structures than mnucfDNA does. Right: Shannon's entropy of uscfDNA is lower than that of mnucfDNA, indicating reduced sequence complexity. Reprinted from [19] under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

2.2. Epigenomic Aberrations and Stability Determinants

In addition to their physical structure, cfDNA and cfRNA fragments carry a deeper layer of information encoded in their epigenome [19,20]. These modifications, which regulate gene expression without altering the DNA sequence itself, are critical hallmarks of cancer and are faithfully preserved on the cfDNA released from tumor cells [20,26]. DNA methylation, the addition of a methyl group to cytosine bases (5mC), is a well-documented event in OSCC, where the hypermethylation of tumor suppressor genes is a key step in carcinogenesis [20,26,27]. These aberrant methylation patterns can be detected in cfDNA from both saliva and plasma, offering a powerful, noninvasive tool for diagnosis and prognosis [20,26].

More recently, 5-hydroxymethylcytosine (5hmC), an oxidized derivative of 5mC generated by TET enzymes, has emerged as another vital epigenetic marker [27]. In healthy tissues, 5hmC is enriched in transcriptionally active regions such as gene bodies and enhancers, playing a key role in gene regulation [27]. A global loss of 5hmC is a recognized feature of many cancers and is often associated with more aggressive tumors and a poorer prognosis [27]. Because 5hmC patterns are

tissue- and cancer specific, their analysis in cfDNA provides a detailed signature that can help identify a tumor's tissue of origin and its malignant state [25,27]. The physical packaging of DNA also contributes to this epigenetic story; nucleosome positioning dictates which regions of the genome are accessible, influencing both fragmentation patterns and gene expression [19,22]. Thus, cfDNA fragment patterns, nucleosome footprints, and methylation/hydroxymethylation signatures are intrinsically linked, creating a multilayered biomarker that reflects both the genetic and epigenetic state of a tumor [19,27].

The ultimate utility of these biomarkers, however, depends entirely on their ability to survive the hostile environment of the oral cavity. Saliva is rich in various enzymes, including nucleases such as DNase I, that can rapidly degrade nucleic acids [17,21,23]. Indeed, studies have shown that unprotected, exogenous DNA or RNA added to saliva is degraded within minutes [21]. However, endogenous salivary cfDNA and cfRNAs are remarkably stable [21,23]. Salivary cfDNA shows a calculated ex vivo half-life of approximately 13–14 hours at room temperature, whereas in vivo cfDNA in blood is commonly estimated to have a half-life of 16 minutes to 2.5 hours [22,23]. This surprising stability is attributed to protective mechanisms that shield cfNAs from enzymatic attack [21,30].

A primary protective strategy is encapsulation within carrier vesicles [18]. Extracellular vesicles, such as exosomes, are shed from cells and act as natural nanocarriers, and their lipid bilayer membranes form a robust shield around their nucleic acid cargo, protecting it from degradation [18,22]. Similarly, cfNAs can associate with macromolecular complexes, such as proteins and lipoproteins, which also provide protection [21,28]. A significant fraction of salivary RNA is associated with such complexes, as it cannot pass through fine-pore filters and is rapidly degraded upon exposure to detergents that disrupt these structures [21]. The majority of extracellular DNA (exDNA; a subset of cfDNA) in saliva appears to be associated with larger structures such as cell debris and apoptotic bodies, as its concentration is dramatically reduced by high-speed centrifugation [24]. This intricate interplay between nuclease activity and protective carrier mechanisms ultimately determines the half-life and integrity of cfNAs in saliva, shaping the quality of the biomarkers we seek to detect [21,24]. The resulting cfDNA molecules that survive this journey represent a rich substrate for diagnostics, carrying a dense, multilayered repository of information that spans their cellular origin, epigenetic state, and physical associations (Figure 5) [19].

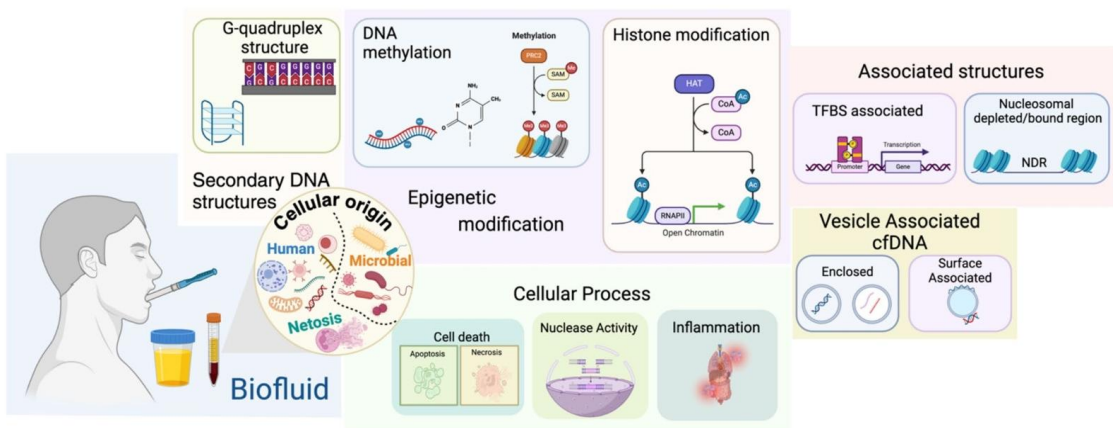


Figure 5. Diverse molecular and cellular information enriched in salivary biofluids. Cell-free DNA (cfDNA) is a multilayered repository of biological data. Analysis can reveal its cellular origin (human, microbial); key epigenetic modifications (e.g., DNA methylation, histone modifications); its physical state, such as being enclosed in vesicles or associated with transcription factors; and the cellular processes (e.g., apoptosis, inflammation) that govern its release into the biofluid. Reprinted from [19] under the terms of the Creative Commons Attribution 4.0 International License.

2.3. Interactions with the Oral Microbiome and External Factors

The oral cavity is a dynamic ecosystem, a dynamic ecosystem where the subtle signals of early-stage disease must be distinguished from a torrent of physiological and environmental noise [31–51]. A primary source of this biological variability is the oral microbiome, a complex community of microorganisms whose composition is in constant flux, profoundly shaped by a host of external factors that act as significant confounders in diagnostic analyses [39,41]. Lifestyle choices, for example, are powerful modulators of this microbial landscape. Dietary patterns, ranging from high-sugar and carbohydrate intake to the consumption of nitrate-rich vegetables, actively select for different bacterial populations; high-sugar diets, for example, encourage the proliferation of acidogenic taxa such as *Streptococcus*, which can lower salivary pH, whereas nitrate consumption can foster a healthier microbiome rich in *Rothia* and *Neisseria* [37,48,50]. Tobacco use represents another potent disruptor, creating an anaerobic environment that depletes health-associated genera such as *Neisseria* while promoting the growth of pathogenic communities, including *Fusobacterium*, thereby inducing a state of dysbiosis that can obscure or mimic disease signatures [35,38,39,41]. Even the simple act of eating can introduce immediate confounding effects, with postmeal increases in plasma lipids potentially creating technical biases that reduce the measured concentration of circulating cell-free DNA (ccfDNA) [47].

In addition to lifestyle, the very biology of the human body introduces a temporal layer of complexity through diurnal and circadian rhythms. The oral environment is not static over a 24-hour period; it ebbs and flows with the body's internal clock [40,46,49]. Compelling evidence has demonstrated that the oral microbiota itself possesses an endogenous circadian rhythm, with the relative abundance of key phyla such as *Proteobacteria* and genera such as *Neisseria*, *Fusobacterium*, and *Streptococcus* oscillating independently of eating or sleeping cycles [40]. This intrinsic rhythmicity is mirrored in the host's inflammatory state and hormonal milieu; proinflammatory cytokines such as IL-1 β and hormones such as cortisol exhibit clear diurnal patterns, typically peaking in the morning before gradually declining throughout the day [36,46,49]. While some studies have not revealed a significant circadian influence on the total concentration of ccfDNA [47], it is clear that the underlying biological landscape, from microbial composition to inflammatory activity, has constant, predictable fluctuations. This rhythmic biological noise poses a formidable challenge, as a sample taken in the morning may present a fundamentally different molecular signature than one taken in the evening from the same individual, complicating the establishment of stable diagnostic baselines.

To distinguish subtle pico-stage OSCC signatures from background biological variability, it is imperative to develop computational tools that can account for and mitigate these multifaceted confounders. Traditional batch effect correction algorithms, such as ComBat, which was designed for genomic data, often prove inadequate for the unique statistical properties of microbiome data, namely, its zero-inflated and overdispersed nature [43,45]. This has spurred the development of adaptive baseline algorithms specifically engineered to address this complexity. Approaches such as Conditional Quantile Regression (ConQuR) represent a significant advancement, employing a nonparametric, two-part model to remove batch effects directly from microbial read counts without making restrictive assumptions about the data's distribution [45]. Similarly, other machine-learning tools, such as Procrustes and pyComBat, have been designed to correct for technical and cross-platform variations in RNA-seq and other high-throughput data [32,43]. These sophisticated frameworks move beyond simple normalization; they learn to disentangle true biological signals of interest from the predictable noise introduced by diet, smoking, and circadian rhythms, creating an adaptive baseline that is crucial for the high sensitivity and specificity required in early cancer detection [42,45,51].

3. Assay Design and Analytical Engineering

3.1. Multiplex Capture Chemistry and Microfluidics

The development of a clinically viable liquid biopsy such as SalivaSense 2.0 hinges on solving a profound engineering challenge: the specific and efficient capture of infinitesimally small quantities of cell-free DNA (cfDNA) and cfRNA from saliva [52–61]. Saliva is a notoriously complex biofluid containing a mixture of genomic DNA from shed epithelial cells, a diverse oral microbiome, and various enzymes, including nucleases that actively degrade the very nucleic acids targeted for analysis [53,55]. the low concentration and fractional abundance of tumor-derived nucleic acids. For example, in blood plasma, the average concentration of cfDNA is approximately 30 ng/mL in healthy individuals, and the tumor-derived fraction (ctDNA) in cancer patients can be as low as 0.01% of this total [54,58,61]. To overcome these obstacles, SalivaSense 2.0 is architected around an automated microfluidic platform. The design of such a platform draws from a diverse toolkit of biochemical techniques, each optimized for the challenging task of isolating rare nucleic acid targets from a complex medium. This approach is fundamental to achieving the precision required for clinical diagnostics, as it minimizes sample volumes, drastically reduces processing times, lowers the risk of cross-contamination, and ensures a standardized, reproducible workflow that is less dependent on operator skill [52,54].

At the heart of the assay's capture technology are superparamagnetic beads, which serve as mobile solid-phase extraction (SPE) media [52]. This strategy offers key advantages over static column-based methods, including a vast surface area for nucleic acid binding and compatibility with high-throughput automation, enabling the simultaneous processing of dozens of samples with minimal hands-on time [56,59]. The surface chemistry of these beads is tailored for maximum recovery. In addition to traditional silica-coated beads that bind nucleic acids under specific chaotropic salt conditions [52,54], SalivaSense 2.0 employs more advanced chemistry using amine-conjugated magnetic beads. These beads are activated by homobifunctional crosslinkers that covalently couple to primary amines on ligands or adapters; native DNA is captured predominantly via electrostatic/chaotrope-mediated interactions unless amine-modified DNA is used [61]. This powerful dual-binding mechanism can complete the capture process in under 10 minutes and has been demonstrated to increase extraction efficiency by more than 50% compared with the QIAamp manual kit [61].

To analyze a comprehensive panel of cancer-specific signatures simultaneously, the assay integrates a highly specific hybrid capture methodology. This is achieved by using pools of thousands of custom-synthesized, 5'-biotinylated single-stranded oligonucleotide baits designed to selectively capture all exons of target genes, specific introns involved in cancer-related fusions, and critical promoter regions [57]. To analyze the epigenome, which provides powerful biomarkers for cancer detection and tissue-of-origin mapping [53,60], the platform is designed to be compatible with enrichment techniques such as immunoprecipitation or target capture with probes following bisulfite treatment [60]. For stable capture of cfRNA, a notoriously fragile molecule, the system can employ established methods such as magnetic beads coated with oligo-dTs to bind eukaryotic mRNA [52].

Rigorous analytical validation confirms the performance of this integrated system. The recovery efficiency, a critical metric for any liquid biopsy, is determined via spike-in experiments where known quantities of synthetic DNA are added to cfDNA-negative plasma [54,56,59]. These validation studies show varied performance across platforms. For example, a manual platform (QIAamp) has demonstrated recovery efficiencies between approximately 52% and 80% depending on the study and input amount [54,56,61]. While some automated platforms perform similarly or even below this range [56], the implementation of optimized chemistries with homobifunctional crosslinkers has been shown to achieve recovery rates exceeding 93%, significantly outperforming the manual kit in a direct comparison [61]. Crucially, the assay demonstrates excellent linearity across a wide dynamic range of inputs, from picogram to nanogram levels, with strong coefficients of determination ($R^2 > 0.95$) [59]. This validated performance ensures that the assay can accurately and reliably quantify the faint

molecular signals of cancer across the full spectrum of clinical presentations, from the earliest pico-stages to advanced disease.

3.2. Ultralow-Input Quantification and Nanopore Barcoding

The journey to detect oral squamous cell carcinoma (OSCC) at its nascent, pico-stage begins with a formidable challenge: the profound scarcity of target molecules in a vast biological background. Cell-free DNA (cfDNA) and RNA extracted from saliva or plasma often yield mere nanograms of material, with the tumor-derived fraction representing a minuscule portion of the total [62–79]. This ultralow-input reality demands a paradigm shift in assay design, moving beyond conventional methods to embrace technologies that can amplify and quantify the faintest of molecular whispers. The narrative of modern diagnostics is thus one of amplification, where single molecules are given a powerful voice, enabling their detection against the noise of a complex sample matrix.

At the forefront are nucleic-acid amplification strategies for ultra-low input detection that bypass the need for cumbersome thermal cycling, making them ideal for point-of-care applications. Droplet digital PCR (ddPCR), which uses thermal cycling, has proven its mettle by partitioning a single sample into thousands of individual microreactors. This partitioning allows for the absolute quantification of rare genetic variants with exceptional sensitivity, successfully detecting TP53 mutations in head and neck cancer patients at fractional abundances as low as 0.01% [71,79]. In addition to PCR, rolling-circle amplification (RCA) offers a versatile isothermal engine that uses a circular DNA template and a highly processive polymerase to generate exceptionally long, repetitive DNA concatemers from a single binding event. This immense amplification factor transforms single-target recognition into a multitude of detectable reporter molecules, pushing detection limits into the picomolar range and providing a robust platform for nanopore-based immunoassays [64,74].

To further refine this ability, CRISPR–Cas systems have been ingeniously repurposed as triggers for amplification. In one such strategy, the Cas9 nickase, guided by specific sgRNAs, creates precise nicks in a DNA target to initiate a strand displacement amplification (SDA) cascade [72,73]. This method, known as CRISDA, leverages the exquisite specificity of Cas9 to launch a true isothermal reaction, achieving attomolar sensitivity and the ability to discriminate single-nucleotide differences without an initial heat denaturation step [73]. These amplification technologies provide the critical first step: turning a handful of molecules into a detectable population. The next challenge is to read this information with single-molecule precision.

This is where the story turns to nanopore sequencing, a technology that provides direct, real-time, electrical readouts of individual nucleic acid molecules [66,67,75]. As a single strand of DNA or RNA is driven through a protein nanopore, it produces a characteristic disruption in an ionic current, a molecular "squiggle" that can be decoded into its base sequence [65,66]. However, the inherently high error rate of this process has historically been a barrier to detecting rare variants. The solution lies in a powerful concept: reading the same original molecule multiple times. The CyclomicsSeq technique masterfully achieves this goal by circularizing and concatenating short cfDNA fragments before sequencing. The resulting long molecule contains many tandem copies of the original template, and by generating a consensus sequence from these copies, the base-calling accuracy is increased by an approximately ~60-fold, enabling the confident detection of mutations at frequencies as low as 0.02% [75]. This principle of error correction, also known as the use of unique molecular identifiers (UMIs) in next-generation sequencing [62], is complemented by the development of optimized nanopore barcoding protocols that allow for high-throughput, multiplexed analysis of numerous ultralow-input samples in a single run [67]. Together, these innovations in amplification and single-molecule consensus sequencing construct a formidable toolkit, transforming the challenge of ultralow-input samples from an insurmountable barrier into a solvable engineering problem.

3.3. Sequencing and Signal-Integration Pipelines

Once the scarce molecular targets from a saliva sample have been amplified into a detectable signal, the subsequent challenge is to accurately sequence this material and integrate the complex

data streams into a coherent, clinically actionable result. This requires a sophisticated fusion of sequencing platforms and analytical pipelines, each contributing a unique strength to the diagnostic narrative. The established workflows of Illumina short-read sequencing, particularly when coupled with targeted bisulfite or oxidative bisulfite sequencing, have long served as the bedrock for methylation analysis [66,68,69]. These methods, while powerful, often rely on harsh chemical treatments such as bisulfite conversion, which can degrade the already limited cfDNA and cannot, on their own, distinguish between the crucial epigenetic marks of 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) [66,69,70]. While modified protocols such as oxidative BS (oxBS-seq) and TET-assisted BS-seq (TAB-seq) were developed to resolve this ambiguity, they introduce additional complexity and potential for DNA damage, underscoring the need for a more direct approach [66,68].

Nanopore sequencing represents a direct approach, offering a revolutionary alternative that reads epigenetic modifications in their native state. As a DNA or RNA molecule traverses a nanopore, modified bases generate unique electrical disturbances, or "squiggles," that differ from their canonical counterparts [66,67]. The translation of this raw signal into a high-resolution epigenome map is a significant computational feat addressed by a growing ecosystem of specialized analytical tools. Pipelines built around software such as Megalodon, Nanopolish, DeepSignal, and Guppy employ advanced algorithms, from hidden Markov models to deep neural networks, to call the methylation status directly from electrical data [65,66]. Systematic benchmarking of these tools has been crucial, revealing top performers such as Megalodon and Nanopolish and paving the way for consensus strategies such as METEORE, which integrates outputs from multiple callers to increase accuracy and reliability [65,66]. This direct, conversion-free analysis not only preserves valuable sample material but also provides long-read information capable of phasing epigenetic marks across entire molecules.

Parallel to the evolution of sequencing, a new frontier is emerging in the form of electrochemical nanobiosensors, which push the limits of sensitivity and bring diagnostics to the point of care [63,64,77]. These devices translate the molecular recognition of a methylated site into a direct, measurable electrical signal. The true power of this technology is unlocked through analytical engineering at the nanoscale. By modifying electrode surfaces with advanced nanomaterials, such as gold nanoparticles (AuNPs), reduced graphene oxide (rGO), or self-assembled DNA tetrahedra, the sensor performance can be dramatically improved [76–78]. These nanostructures create a vastly increased surface area for capturing target molecules and enhance electronic signal transduction, enabling synergistic amplification schemes that can achieve attomolar detection limits [78]. A key advantage of this platform is its capacity for a real-time, locus-aggregate methylation readout. Unlike sequencing, which requires extensive postacquisition data processing, an electrochemical biosensor provides a direct quantitative readout as a binding event occurs. For example, a sensor can be functionalized with methylation-specific antibodies conjugated to a nanocomposite; the binding of methylated cfDNA directly triggers a change in current, providing a real-time measure of the target's concentration [76]. This integration of signal generation and detection into a single step is consistent with point-of-care use. The ultimate vision for SalivaSense 2.0 lies in the intelligent integration of these pipelines: leveraging the broad discovery power of Illumina and nanopore sequencing for comprehensive profiling while deploying electrochemical nanobiosensors for rapid, ultrasensitive, and cost-effective quantification of key validated biomarkers. This multimodal approach, where diverse data streams converge, is the cornerstone of building the robust translational pipelines necessary to achieve pico-stage oral cancer detection from the laboratory to the clinic.

4. AI-Enhanced Bioinformatics and Machine-Learning Pipelines

4.1. Preprocessing, Error Suppression and Methylation Calling

The journey from a raw saliva sample to a clinically actionable biomarker is paved with computational challenges that demand a sophisticated bioinformatic pipeline. The initial cfDNA and cfRNA sequences obtained from high-throughput technologies are not pristine signals but are instead

fraught with noise, missing data, and systematic errors that can obscure the subtle epigenetic signatures of early-stage cancer [80–92]. Therefore, the initial preprocessing and error suppression phase is not merely a procedural step; it is the foundational bedrock upon which the entire diagnostic framework is built. Traditional, simplistic methods for handling missing data, such as filling gaps with the mean or median value of a feature, risk distorting the underlying biological distributions and introducing noise that can mislead downstream analyses [88]. To navigate this, advanced frameworks employ generative models such as generative adversarial imputation networks (GAINs), which transcend simple substitution by learning the complex, multidimensional relationships within the omic data. By training a generator to produce realistic data points and a discriminator to distinguish them from real data, these networks can impute missing values that are consistent with the learned feature distributions, thereby preserving the integrity of the dataset [88].

In addition to missing data, the most formidable challenge in cfDNA analysis is the overwhelming background noise from sequencing artifacts, which can easily eliminate the subtle signals of circulating tumor DNA (ctDNA) [91]. A groundbreaking strategy to overcome this problem is integrated digital error suppression (iDES), a dual-pronged approach that combines molecular and computational techniques to achieve unprecedented sensitivity [91]. The first prong is molecular barcoding, where unique molecular identifiers (UMIs) are attached to each original DNA fragment before amplification. This allows every resulting sequence read to be traced back to its parent molecule, enabling the computational distinction between true biological variants and errors introduced during PCR or sequencing [91]. The second prong is *in silico* "background polishing." This involves creating a detailed error profile from a cohort of healthy individuals to model the highly recurrent, stereotypical background noise, often caused by oxidative damage during sample preparation, that plagues sequencing data. This background model is then used to computationally "polish" new samples by statistically removing artifacts that are indistinguishable from this established noise profile [91].

Crucially, these two methods are synergistic. Molecular barcoding effectively eliminates random, stochastic errors while background polishing targets and eliminates the most common, position-specific artifacts. When combined, they achieve a remarkable ~15-fold reduction in error rates, increasing the fraction of error-free genomic positions from ~90% to ~98% [91]. This dramatic increase in the signal-to-noise ratio pushes the limits of detection, enabling the identification of ctDNA at allelic fractions as low as 4 in 100,000 cfDNA molecules [91]. Once these high-fidelity data are generated, methylation calling is performed by calculating per-read calls for each CpG site, which quantify the level of methylation and serve as the core input for subsequent machine learning models [81,90]. This meticulous pipeline of imputation, error suppression, and methylation calling transforms the raw, chaotic output of the sequencer into a clean, quantitatively precise dataset ready for the discovery of robust epigenetic signatures.

4.2. Multi-Omic Feature Selection and Integration

With a foundation of high-quality, denoised data, the next critical stage involves navigating the vast, high-dimensional landscape of the multiomic epigenome. The challenge shifts from cleaning the data to extracting meaning from it: selecting the most informative features and integrating them in a way that captures the intricate, synergistic interplay between different molecular layers. A brute-force approach, in which tens of thousands of features are fed into a model, is computationally intractable and risks "the curse of dimensionality," where models learn noise instead of true biological signals [88,92]. To counter this, a robust feature selection process is essential. While simple variance filters are common, they can be unstable and biased by outliers [88]. A more rigorous method is robust feature selection (RFS), which employs bootstrap sampling, repeatedly creating subsets of the data, to identify features that are consistently ranked as important across these different perturbations. This ensures that the selected feature set is not an artifact of a specific data split but is truly stable and biologically relevant [88]. Another powerful technique is the chi-square test, which

can identify features that are statistically significantly associated with clinical outcomes, such as survival status [85].

Once a stable feature set is identified, the challenge becomes one of integration. The simplest strategy, known as **early fusion**, involves concatenating features from different omics (e.g., cfDNA methylation, cfRNA expression) into a single, long vector. However, this approach treats all features equally and often fails to capture the unique characteristics and hierarchical relationships between different data types [88,89]. At the other end of the spectrum is **late fusion**, where separate models are built for each omic type and their final predictions are combined [80,89]. This preserves modality-specific patterns but may miss crucial cross-omic interactions. The frontier of innovation lies in intermediate fusion strategies that model the data's inherent structure, with **graph neural networks (GNNs)** emerging as a particularly powerful paradigm [80]. GNNs represent multiomic data as a biological network, where genes and other molecules are nodes and their interactions are edges. This approach explicitly models both intraomic (e.g., gene-gene) and interomic (e.g., miRNA-gene) connections, allowing models such as graph convolutional networks (GCNs) and graph attention networks (GATs) to learn from the biological context of the data [85,87]. GATs, in particular, can dynamically weigh the importance of different molecular interactions, offering a more nuanced feature extraction that has proven superior for complex classification tasks such as determining cancer molecular subtypes [87].

Pushing this frontier even further are **transformer-based architectures**, inspired by their revolutionary success in natural language processing [89,92]. Models such as DeePathNet leverage the transformer's self-attention mechanism, which allows the model to dynamically assess the importance of each feature in the context of all others, capturing complex, nonlinear dependencies [92]. A key innovation of DeePathNet is its integration of domain-specific biological knowledge. Instead of learning from individual genes, it first encodes multiomic features into 241 literature-curated cancer pathways. The transformer then learns the interactions *between these pathways* [92]. This hierarchical approach is not only highly accurate but also inherently explainable. By analyzing the model's internal workings, it is possible to create **transformer-based explainable dashboards** that highlight which specific pathways drive a prediction for an individual patient. This provides unprecedented interpretability, aligning model predictions with known biological processes and TCGA profiles, thereby moving beyond a "black box" prediction to deliver actionable, pathway-level biomarker insights [92]. To navigate the vast landscape of potential model architectures and hyperparameters, automated machine learning (**AutoML**) platforms can be employed. Tools such as H2O AutoML systematically train and evaluate a diverse array of models, from gradient boosting machines to deep neural networks and stacked ensembles, and rank them on a "leaderboard," streamlining the process of identifying the optimal algorithm for a given dataset and task [82].

4.3. Cross-Cohort Harmonization and Federated Transfer Learning

The journey to develop a universally effective diagnostic tool from complex omics data begins by confronting two fundamental hurdles: data are fragmented across numerous institutions, and they are overwhelmingly derived from individuals of European ancestry [93–103]. This reality creates significant "cohort bias", nonbiological variations arising from different scanners, lab preparations, and acquisition protocols, which can dramatically skew results and lead to models that fail when applied to new populations [94,95]. The poor transferability of predictive scores developed in one ancestral group to another, particularly from European to African ancestries, represents a critical barrier to equitable precision medicine, which threatens to exacerbate existing health disparities [100,103]. To dismantle these barriers, a sophisticated pipeline combining data harmonization with privacy-preserving federated learning is not only advantageous but also essential for building robust, globally deployable models.

The first crucial step is data harmonization, a computational process that retrospectively aligns datasets from diverse sources to create a cohesive and analyzable whole [94]. This involves computationally adjusting for technical variability, ensuring that biological signals are not obscured

by noise from different equipment or protocols [94,95]. With data made comparable, federated learning (FL) offers a transformative solution to the challenge of data silos. This decentralized paradigm allows multiple institutions to collaboratively train a single, powerful AI model without ever sharing raw, sensitive patient data [93,96,99]. Instead of pooling data, each institution trains the model locally and shares only the resulting model parameters, which are then aggregated to create an improved global model [93,99]. This approach has proven remarkably effective, with studies on large-scale genomic data demonstrating that federated models can achieve predictive performance nearly identical to that of centralized models trained on pooled data, all while preserving patient privacy [99].

To address ancestral disparities specifically and ensure global utility, the pipeline is elevated with federated transfer learning (FTL) [96]. This advanced technique intelligently leverages the wealth of data from well-represented populations to increase predictive accuracy in underrepresented groups [100]. By identifying and utilizing the genetic similarity shared between, for instance, European and African populations, FTL can adapt genetic effect sizes from the data-rich cohort to the target cohort, significantly improving the model's predictive power for non-European individuals [100,101]. For true global deployment where trust is paramount, the entire process is fortified with state-of-the-art privacy-preserving analytics. While standard FL prevents direct data sharing, sensitive information can still potentially be inferred from model updates [93,95]. Therefore, advanced cryptographic methods such as multiparty homomorphic encryption, which enables direct computation on encrypted data, and perturbation techniques such as differential privacy, which adds statistical noise to mask individual contributions, are integrated to provide end-to-end security without compromising model accuracy [93,98]. Finally, the powerful, harmonized insights generated from these secure federated systems can fuel the creation of open-data commons and large-scale, integrative biomedical knowledge graphs, creating a virtuous cycle where securely shared knowledge accelerates new discoveries for all populations [96].

5. Validation Frameworks

5.1. Analytical benchmarks

Before SalivaSense 2.0 can be deployed in a clinical setting, its fundamental performance characteristics must be meticulously defined. This foundational stage of validation is guided by established principles, such as those outlined in the Minimum Information for Publication of Quantitative Digital PCR Experiments (dMIQE) guidelines, which provide a crucial roadmap for ensuring accuracy, reproducibility, and transparency [104–113]. The initial challenge is to determine the absolute boundaries of the assay's capabilities, including its limit of detection (LOD), linearity, and reproducibility, which are essential for guaranteeing that the technology can reliably quantify the minute amounts of cfDNA and cfRNA characteristic of early-stage disease.

These benchmarks are established in a controlled environment, and well-characterized reference standards are used to create a pristine baseline. A validation workflow, for example, can employ commercially available cfDNA reference standards spiked into a clean matrix such as DNA-free plasma to systematically evaluate performance across a range of concentrations and sample input volumes [113]. This process allows for the precise determination of linearity, where a strong linear correlation (e.g., $R^2 \geq 0.99$) between the known input amount and the measured output confirms that the assay performs consistently and quantitatively across its dynamic range [113]. The use of certified reference materials (CRMs), such as the NIST Standard Reference Material® 2372a, represents the gold standard in this process. These materials provide genomic DNA with certified values for both mass concentration (e.g., ng/μL) and copy number concentration, which are metrologically traceable to the International System of Units (SI) [111]. By calibrating assays against these CRMs, it becomes possible to ensure not only precision but also trueness, allowing for the harmonization of results across different laboratories and platforms, a critical step for any diagnostic test intended for widespread use [111,112].

However, achieving linearity and a low LOD in a single laboratory is insufficient. True analytical robustness is demonstrated through reproducibility, a measure of how consistent the results are when variables are introduced. Interlaboratory and interoperator variability are significant hurdles in biomarker validation, often stemming from subtle differences in protocol execution [110]. To combat this, validation frameworks must incorporate strategies to minimize human error and standardize procedures. The implementation of standardized operating procedures (SOPs) and the use of automated systems, such as robotic liquid handlers for RNA isolation and PCR setup, have been shown to dramatically reduce operator-induced variance and improve the correlation of results between different facilities, with correlation coefficients as high as 0.96 for some mRNA markers [110]. Furthermore, conducting reproducibility studies where multiple independent operators perform the same extraction and analysis protocols on identical reference materials is a key validation step. Consistent yields, fragment size profiles, and cfDNA percentages between operators confirm the method's robustness and readiness for broader application [113].

Once these core benchmarks are established, the assay must be challenged with interferent stress tests designed to mimic the complex, real-world conditions of the oral cavity. The mouth is not a sterile test tube; it is a dynamic environment teeming with substances that can inhibit PCRs, degrade nucleic acids, or contaminate the sample with nontumorous genetic material. One of the most significant challenges is exposure to tobacco tar. Cigarette smoke is a complex aerosol containing more than 9,500 chemical compounds, including a host of reactive oxygen species (ROS), carcinogenic organic compounds, and heavy metals such as cadmium [106,107]. These components can inflict direct oxidative damage on DNA and RNA while also potentially inhibiting the enzymatic reactions central to PCR-based assays, thereby compromising the integrity of the results [106]. This intricate web of molecular damage, illustrated in (Figure 6) [106], underscores the necessity of stress-testing the SalivaSense 2.0 assay against tobacco exposure. The figure details how free radicals from smoke not only inflict direct DNA damage but also trigger inflammatory pathways and lipid peroxidation, all of which could introduce confounding signals that must be distinguished from the signature of OSCC.

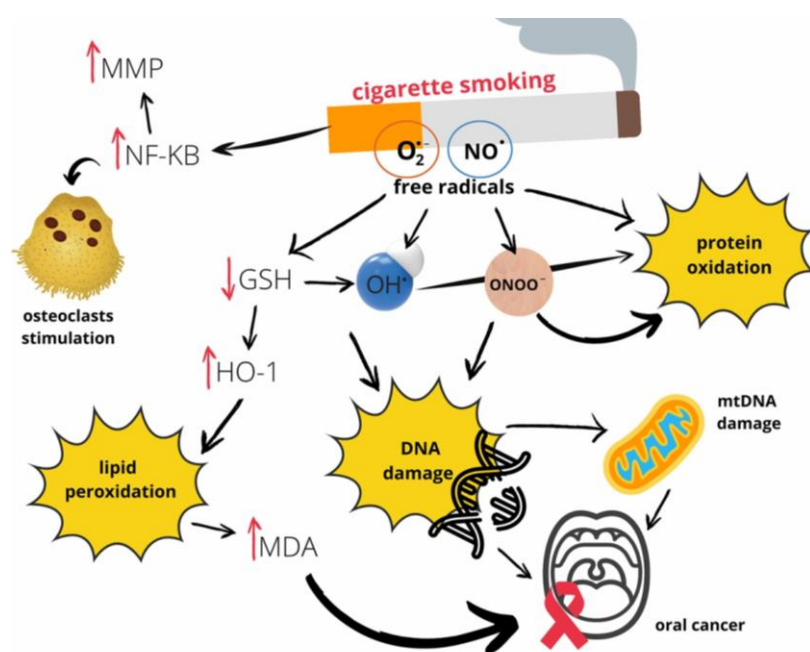


Figure 6. Molecular Cascade of Tobacco-Induced Oral Damage. Cigarette smoke unleashes free radicals (e.g., the superoxide anion radical O_2^- and nitric oxide NO^-) that trigger a cascade of cellular damage. This includes direct DNA and mitochondrial DNA (mtDNA) damage, protein oxidation, and lipid peroxidation (measured by markers such as malondialdehyde (MDA)), ultimately contributing to oral cancer. The pathway also involves the activation of inflammatory signals such as NF- κ B, which upregulates matrix metalloproteinases (MMPs)

involved in tissue degradation. Reprinted from [106] under the terms of the Creative Commons Attribution 4.0 International License.

Similarly, exposure to mouthwash and alcohol introduces another layer of complexity. Ethanol itself can disrupt cellular homeostasis, but its primary metabolite, acetaldehyde, is classified as a Group 1 carcinogen that readily forms DNA adducts [105,106]. Critically, rinsing the mouth with an alcoholic beverage can cause a rapid, dramatic spike in salivary acetaldehyde to levels exceeding 1,000 μM , far above the $\sim 100\text{ }\mu\text{M}$ threshold considered mutagenic, even without swallowing [105]. This short-term, high-concentration exposure from the beverage itself, independent of systemic metabolism, represents a potent interferent that can directly compromise the cfDNA and cfRNA targets that SalivaSense 2.0 is designed to detect. The mechanisms through which alcohol exerts this damaging effect are multifaceted and involve the generation of free radicals that drive protein oxidation, lipid peroxidation, and direct DNA damage, as depicted in (Figure 7) [106]. These pathways represent a significant source of biological noise that the assay must be robust enough to overcome.

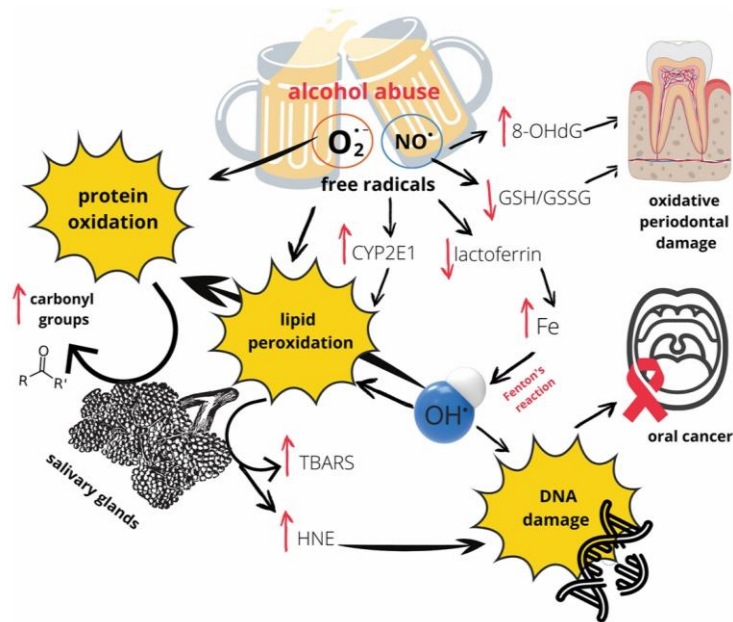


Figure 7. Mechanisms of alcohol-induced oral carcinogenesis. Alcohol abuse generates free radicals that initiate widespread oxidative stress, leading to lipid peroxidation, protein oxidation, and direct DNA damage. The figure highlights the role of enzymes such as CYP2E1 in the metabolism of ethanol and the Fenton reaction, where iron (Fe) catalyzes the formation of highly reactive hydroxyl radicals ($\text{OH}^\bullet$). These insults contribute to both oral cancer and oxidative periodontal damage, which are potential sources of confounding cfDNA in salivary diagnostics. Reprinted from [106] under the terms of the Creative Commons Attribution 4.0 International License.

Finally, the validation framework must account for biological interferents inherent to the oral cavity, most notably gingival crevicular fluid (GCF). GCF is a serum transudate that continuously flows from gingival tissues into the oral cavity, and its volume increases significantly with periodontal inflammation [104]. This fluid is a rich source of host-derived cfDNA released from apoptotic and necrotic periodontal cells. Studies have shown that cfDNA concentrations in GCF are orders of magnitude higher than those in saliva or plasma and are positively correlated with the severity of periodontal disease [104]. This influx of noncancer, inflammation-associated cfDNA presents a formidable challenge, as it can dilute the tumor-specific signal or, worse, generate background noise that could lead to false-positive results. Indeed, common comorbidities such as periodontitis are a primary source of this biological noise, resulting in a distinct and measurable profile of molecular biomarkers that can confound cancer-specific signals (Figure 8) [18].

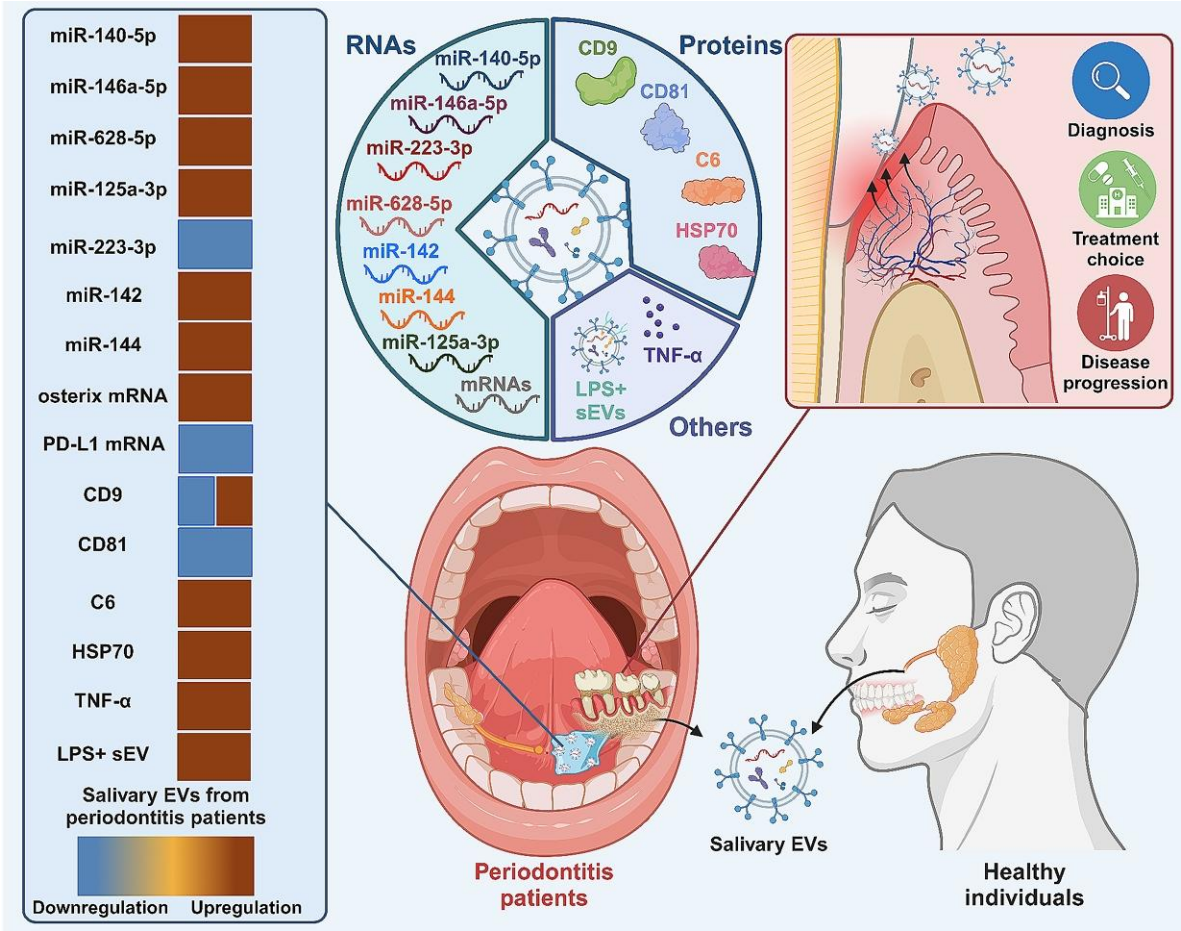


Figure 8: Periodontitis as a significant confounding factor in salivary diagnostics. Comorbidities such as periodontitis introduce significant biological noise. Salivary EVs from periodontitis patients contain distinct, dysregulated molecular signatures, including a panel of upregulated and downregulated RNAs and proteins (e.g., miR-146a-5p and TNF-α). These inflammation-specific markers can obscure or mimic cancer signals, necessitating diagnostic assays with high specificity to differentiate disease states. Reprinted from [18] under the terms of the Creative Commons Attribution 4.0 International License.

Therefore, a pivotal part of the validation for SalivaSense 2.0 involves demonstrating that its multiplex signatures are specific enough to distinguish the epigenetic and transcriptomic fingerprints of OSCC from the considerable background of host cfDNA introduced by common comorbidities such as periodontitis. Successfully navigating these stress tests is what will ultimately prove the technology's clinical utility, transforming it from a promising concept into a life-saving diagnostic tool.

5.2. Clinical Validation Study Designs

Before a new diagnostic can be entrusted with patient lives, it must prove its mettle against the established "gold standard", in the case of oral cancer, the definitive histopathological examination of a tissue biopsy [114–123]. The validation framework for SalivaSense 2.0 is envisioned as a multiact drama, beginning with controlled, retrospective inquiries and culminating in real-world prospective trials that simulate its ultimate clinical application.

The initial act of this validation story often unfolds through **retrospective case-control studies**, which provide a powerful and efficient means of initial assessment. In this design, archived saliva samples from individuals with confirmed oral squamous cell carcinoma (OSCC) are compared against those from noncancerous groups, including healthy individuals and patients with oral potentially malignant disorders (OPMDs) [114,122]. A crucial element here is the establishment of

orthogonal concordance, where the novel biomarker's performance is cross-validated against an established technology, such as comparing a rapid, point-of-care test to a laboratory-based ELISA on the same cohort of samples [114]. This dual-assay approach ensures that the biomarker signal is robust and not an artifact of a single detection method. To build unwavering confidence, these studies must incorporate blinding, where laboratory analysts are unaware of the clinical diagnosis of the samples they are processing. This safeguard is essential to prevent operator bias from influencing the results [115]. The narrative is further strengthened by expanding the investigation to include **blinded external cohorts**, where samples are collected and analyzed by different clinicians and technicians at multiple, independent institutions. Successfully replicating a test's high sensitivity and specificity across different sites and operators demonstrates that its performance is not confined to a single specialized center but is generalizable and reproducible in diverse clinical settings [115].

The final act of clinical validation transitions from looking back at past cases to looking forward to **prospective screening trials**. These studies represent the ultimate test of a biomarker's utility, moving it into the real-world setting where it is intended to be used [116]. Such trials are strategically designed to enroll high-risk populations, individuals with a history of tobacco and/or alcohol use, or betel quid chewing, who stand to benefit most from early detection [119,121]. In this framework, participants are screened with SalivaSense 2.0 at regular intervals, and every suspicious finding is referred to as the definitive histopathology gold standard [115,119]. This design not only confirms the biomarker's ability to identify early-stage cancers and precancerous lesions in a real-world screening context but also allows for the assessment of crucial performance metrics such as the positive predictive value (PPV), the probability that a positive test result correctly identifies the presence of disease [114]. While these prospective trials are resource intensive and require long-term follow-up, they are the indispensable final step to prove that a biomarker can truly change clinical outcomes, reduce mortality, and earn its place in the standard of care [119,121].

5.3. Health-Economic Modeling and Decision-Curve Analysis

A clinically validated test is only truly translational if it offers clear value to patients, clinicians, and the healthcare system. Demonstrating this value requires moving beyond traditional metrics of sensitivity and specificity to the sophisticated realms of health-economic modeling and decision curve analysis (DCA). This stage of the validation journey answers the critical questions: "Is SalivaSense 2.0 a wise investment for our health resources?" and "Will it actually lead to better clinical decisions?"

To address the first question, cost-utility analyses are performed via powerful simulation tools such as Markov models [119,120]. From a societal or healthcare system perspective, these models project the long-term consequences of implementing a new screening program over a patient's lifetime [121]. By integrating data on disease progression, treatment costs, and quality of life, they estimate outcomes in terms of quality-adjusted life years (QALYs) gained and calculate the incremental cost-effectiveness ratio (ICER), the additional cost for every QALY gained [120,121]. This allows for a direct comparison of value for money. These models are not monolithic; they are used to conduct detailed scenario analyses, exploring different implementation strategies to find the optimal path forward. For example, models can compare mass screening of the general population versus targeted screening of high-risk groups or evaluate the cost-effectiveness of different screening intervals (e.g., every 3, 5, or 10 years) [121]. Furthermore, these frameworks can generate crucial equity impact projections and budget impact analyses, estimate the financial consequences of a nationwide rollout and ensure that the proposed strategy is not only cost-effective but also affordable and accessible, particularly in resource-constrained settings [119,121].

Complementing economic evaluations, decision curve analysis (DCA) offers a revolutionary way to assess a biomarker's clinical usefulness from the decision maker's perspective [117]. DCA quantifies the "net benefit" of using a test by weighing the benefit of detecting a true cancer against the harm of an unnecessary biopsy for a false positive [118]. This calculation is performed across a range of "threshold probabilities," which represent how a clinician or patient might weigh those risks

and benefits [117]. The result is a simple curve that visually demonstrates whether using the test to guide decisions is superior to default strategies such as "biopsy all suspicious lesions" or "biopsy none" [117,118]. A test that shows a greater net benefit across a wide range of reasonable clinical preferences is one that can be confidently integrated into practice. To accelerate the collection of evidence needed for both robust clinical validation and these complex economic models, the field is embracing innovative clinical trial designs. Adaptive designs for diagnostics, for example, represent a paradigm shift. These patient-centric designs can evaluate multiple therapies or diagnostics across various biomarker-defined subgroups simultaneously, allowing for modifications as the trial progresses [116]. This flexibility enables researchers to stop futile arms early and focus resources on promising avenues, thus facilitating the rapid evidence generation required to bring transformative technologies such as SalivaSense 2.0 to the patients who need them most quickly than ever before [116].

6. Translational Pipeline and Implementation Science

6.1. Point-of-Care Device Architectures and Sample Logistics

The journey of diagnostic technology from a centralized laboratory to a point-of-care (POC) setting is a tale of miniaturization, simplification, and ingenuity [124–139]. For a platform such as SalivaSense 2.0 to achieve its potential in early cancer detection, it must escape the confines of the traditional lab, shedding its reliance on bulky equipment and highly trained personnel [131]. The key to this transformation lies in sophisticated yet accessible device architectures that place powerful analytical capabilities directly into the hands of clinicians and patients [137]. At the heart of this revolution is the microfluidic "lab-on-a-chip," a device that manipulates minute volumes of fluid within microchannels to perform complex tasks such as sample preparation, amplification, and detection [124,131]. These platforms, often fabricated from low-cost materials such as paper or polymers via techniques such as 3D printing, offer a cascade of advantages: they require minimal samples, shorten processing times, increase maneuverability, and reduce costs, making them ideal for onsite diagnostics [124,130,139].

A significant hurdle for any nucleic acid-based POC test is the need for precise temperature control, as amplification methods such as LAMP or RPA require stable, elevated temperatures [131]. The traditional solution, an external electrical heater, compromises the instrument-free ideal of POC diagnostics [131]. A truly innovative pathway forward borrows a concept from an unexpected source: smart food packaging [131]. Commercially available self-heating methods can use controlled exothermic reactions, such as the hydration of calcium oxide, to generate heat without any external power [131]. By integrating these chemical heaters with phase change materials (PCMs) that absorb and store thermal energy, it is possible to create a disposable, electricity-free heating module capable of approximating the target temperatures needed for isothermal amplification; feedback control or calibration may be required for precise operation [131]. This self-contained thermal management system represents a critical step toward a fully autonomous diagnostic cartridge [131].

The final piece of the POC architecture is the reader, and the modern smartphone has emerged as a surprisingly powerful and ubiquitous tool for this purpose [137]. Equipped with high-resolution cameras, substantial processing power, and wireless connectivity, smartphones can be transformed into sophisticated diagnostic readers with the help of simple, low-cost attachments [137]. For optical biosensors, the phone camera can capture colorimetric or fluorescent signals, which are then quantified by a custom app to determine biomarker concentrations [130,137]. Similarly, in electrochemical systems, a portable detector can interface with a smartphone via Bluetooth to process signals and display results in real time [137]. This integration not only makes the system portable but also connects it to the cloud, enabling data sharing, remote analysis, and the application of machine learning algorithms to increase diagnostic accuracy [127,137].

Beyond the device itself, the logistics of sample handling must be reimaged for remote and resource-limited settings. Saliva is notoriously complex, and its enzymatic components can quickly

degrade target biomarkers such as cfDNA and cfRNA [129,130]. While preservatives are an option, an elegant solution for remote shipping is the use of dried saliva spots (DSSs) [129]. In this method, a saliva sample is applied to filter paper and allowed to dry, a process that stabilizes nucleic acids and proteins by suppressing enzymatic degradation [125,129]. These dried spots can remain stable for up to a month at various temperatures, eliminating the need for cold-chain transport and allowing samples to be safely mailed from a patient's home to a central laboratory for analysis [129]. This simple yet robust technique democratizes access, ensuring that a patient's geographical location is no longer a barrier to receiving advanced molecular diagnostics [129,131].

6.2. Training, Adoption, and Equity Roadmap

The successful translation of SalivaSense 2.0 from a technological marvel to a clinical standard hinges on a carefully orchestrated implementation strategy that prioritizes user training, equitable adoption, and community trust [133,136]. A novel diagnostic, regardless of its precision, will fail if clinicians are not confident in its use, if patients do not trust its purpose, or if systemic barriers prevent its reach into underserved communities [138]. Therefore, a robust roadmap must address the human elements of adoption with the same rigor applied to the device's technical design. A critical first step is ensuring standardized and high-quality sample collection, which can be achieved through innovative training modalities such as augmented reality (AR) [128,132]. AR-guided tutorials can overlay digital instructions onto a user's view of the real world, providing an immersive, step-by-step guide to the saliva collection process [128]. By enabling an expert to offer spatially aligned visual guidance from the learner's point-of-view, AR can simplify complex procedures and ensure that both clinicians and patients can collect samples correctly, a foundational requirement for accurate results [132].

Beyond initial sampling, widespread adoption requires a concerted effort to upskill the broader clinical workforce [133]. Many health practitioners report limited training and confidence in specialized fields, a gap that can be effectively addressed through large-scale, accessible e-learning programs [133]. A national training initiative that provides government-funded, evidence-based modules on the principles of salivary liquid biopsy and the clinical utility of SalivaSense 2.0 would be essential for improving clinician knowledge, building confidence, and attracting new providers to the field [133]. This educational outreach must extend into the community through creative and culturally resonant engagement strategies that build trust where it has been historically eroded [135,136]. Programs that embed health education within welcoming, nonclinical community activities can lower emotional barriers, foster open dialog, and humanize public health messaging around sensitive topics such as cancer [135]. By meeting communities where they are, both geographically and culturally, such initiatives transform outreach from a transactional moment into a trust-building experience [135].

Finally, this entire pipeline must be built on a foundation of equity [138]. This requires moving beyond a one-size-fits-all approach to consent and result delivery and instead developing pathways that are culturally tailored and codesigned with the communities they are meant to serve [134,136]. For genomic testing, informed consent is a complex process, and materials must be accessible to individuals with varying levels of health and digital literacy [134]. Dynamic consent platforms (DCPs) offer a promising digital solution, providing layered information through multiple modalities, but concerns remain about their accessibility for vulnerable populations who may lack internet access or digital literacy [134]. To ensure equity, both digital and nondigital consent options must be available, supported by patient navigators, who can help address language, financial, and cultural barriers [134,136]. Ultimately, a successful translational pipeline ensures not only that the technology works but also that it works for everyone, fulfilling the promise of precision medicine by making advanced diagnostics accessible to all populations [138].

6.3. Regulatory and Data-Governance Pathways

The journey of a novel diagnostic such as SalivaSense 2.0 from a validated concept to a clinical tool is a trek through a complex and evolving international regulatory landscape [140–155]. Successfully navigating this requires a multifaceted strategy that addresses distinct approval pathways, stringent data security mandates, and profound ethical considerations regarding patient data.

In the United States, a promising route is the Food and Drug Administration's (FDA) Breakthrough Devices Program, a voluntary pathway designed to expedite the development, assessment, and review of medical devices that offer more effective diagnosis or treatment of life-threatening or irreversible debilitating conditions [154]. This program provides manufacturers with more frequent interactions with FDA experts and prioritized review, accelerating patient access to innovative technologies [154]. Across the Atlantic, the European Union (EU) presents a dual challenge with the In Vitro Diagnostic Medical Devices Regulation (IVDR) and the new AI Act [149,150]. The IVDR establishes a stringent, risk-based framework for devices used to examine human samples, such as genetic screening tools [150]. Layered on top, the AI Act classifies AI-enabled medical devices as "high risk," imposing demanding obligations for risk management, data governance, transparency, and human oversight [149,150]. This creates a dual conformity assessment that can be particularly burdensome for SMEs [150]. For global health applications, especially in low- and middle-income countries (LMICs) with less developed domestic regulatory frameworks, the World Health Organization's (WHO) Prequalification of In Vitro Diagnostics (PQDx) program serves as a critical gateway, ensuring access to safe, appropriate, and affordable diagnostics [140]. Underpinning all these pathways is the nonnegotiable requirement for robust cybersecurity. Both the FDA and the EU AI Act mandate that cybersecurity is integral to device safety, requiring a "Secure Product Development Framework" (SPDF) and resilience against threats that could compromise device function and lead to patient harm [149,151].

Beyond the device itself, the governance of patient data presents an equally complex frontier. Traditional "broad consent" models are being challenged by more patient-centric approaches such as "dynamic consent" [142]. Implemented via digital platforms, dynamic consent aims to create an ongoing partnership, allowing participants to make granular, study-specific decisions about how their data are used over time [141,142]. While this model enhances participant control and transparency, its implementation faces challenges, including low uptake rates in some studies and the risk of "consent fatigue" [141,142]. A further challenge arises from the powerful capabilities of next-generation sequencing: the potential for incidental discovery of pathogenic germline variants (PGVs) unrelated to the primary diagnostic goal [144]. The management of these secondary findings, which may reveal predispositions to hereditary cancers or other serious conditions, requires a carefully structured pipeline [143,144]. Although some assays filter out most germline alterations, they may still report clinically significant variants in genes such as *BRCA1* and *BRCA2* [152]. This underscores the critical need for clear protocols that emphasize that such tests are not a replacement for dedicated germline testing and mandate that suspected PGVs trigger a referral for confirmatory testing and professional genetic counseling [144,152].

7. Ethical, Social and Policy Considerations

The revolutionary potential of SalivaSense 2.0, a promising technology for detecting oral squamous cell carcinoma (OSCC) at its nascent, pico-stage, extends far beyond the laboratory bench. As we stand on the cusp of integrating such powerful diagnostic tools into clinical practice, we must navigate a complex terrain of analytical hurdles, data governance challenges, and profound social responsibility to ensure equitable access and benefit. The journey from a promising multiplex signature to a globally trusted diagnostic requires deliberate and thoughtful engagement with the ethical, social, and policy dimensions of genomic medicine, especially as the potential applications

for salivary biomarkers extend far beyond a single disease to encompass a wide array of systemic conditions (Figure 9) [18].

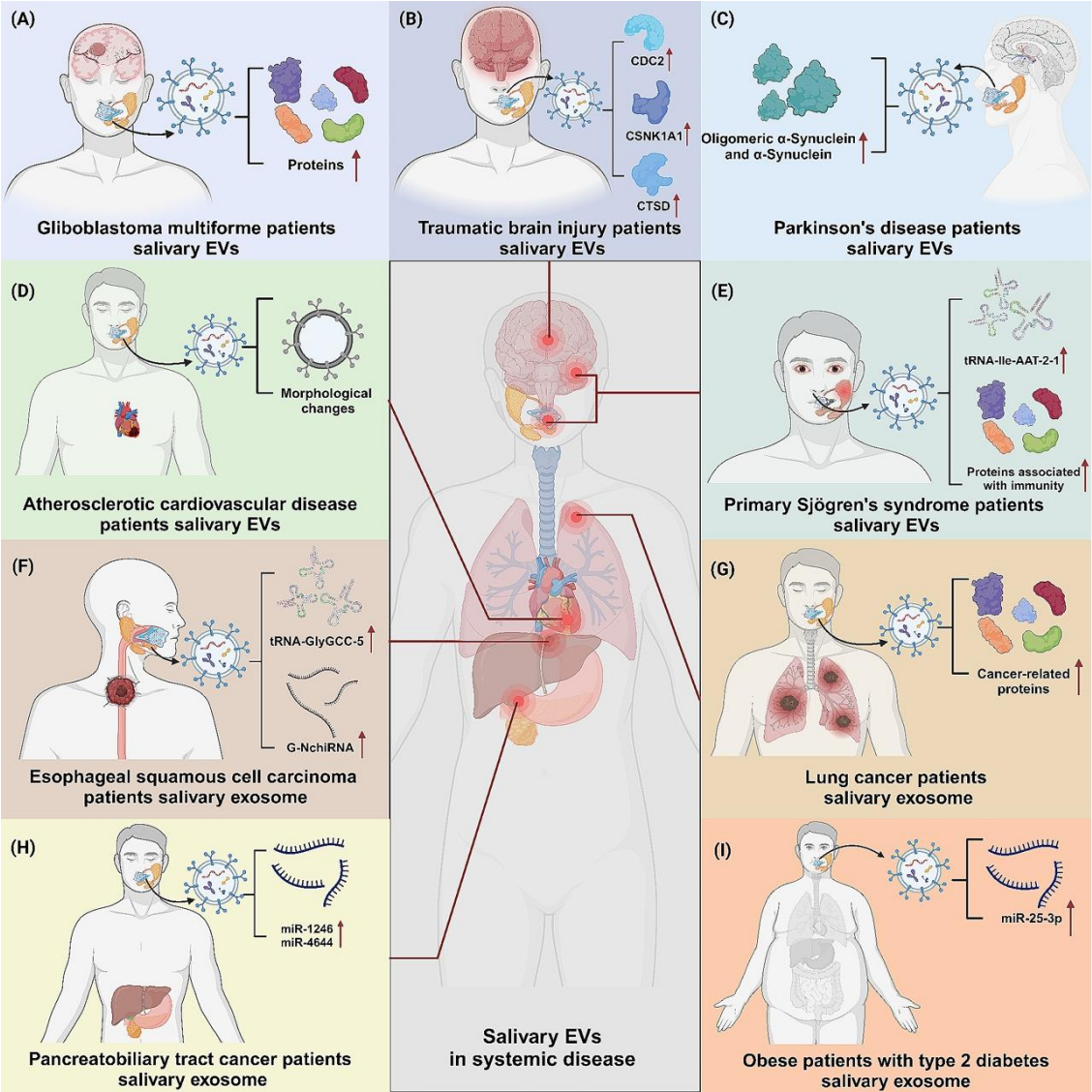


Figure 9. Broad diagnostic potential of salivary extracellular vesicles (EVs) in systemic disease. Saliva serves as a diagnostic window into systemic health. Salivary EVs have been shown to serve as biomarkers for a wide range of conditions beyond the oral cavity, including (A) glioblastoma, (B) traumatic brain injury, (C) Parkinson's disease, (D) atherosclerotic cardiovascular disease, (E) primary Sjögren's syndrome, (F) esophageal squamous cell carcinoma, (G) lung cancer, (H) pancreatobiliary tract cancer, and (I) obesity with type 2 diabetes. Reprinted from [18] under the terms of the Creative Commons Attribution 4.0 International License.

7.1. Analytical Pitfalls and Biological Noise Mitigation

The promise of detecting OSCC at the pico-stage via noninvasive saliva samples is a monumental leap forward, yet its realization hinges on overcoming a formidable analytical challenge: discerning the ultra-low-abundance signal of early cancer from a cacophony of biological and technical noise [156–172]. Liquid biopsy analytes, particularly cell-free DNA (cfDNA) and cell-free RNA (cfRNA), are present in vanishingly small quantities in early-stage disease, allowing their detection akin to finding a needle in a haystack [158,169,170]. This challenge is compounded by the fact that the vast majority of cfDNA in a saliva or blood sample originates from the turnover of healthy cells, creating a significant background of biological noise that can easily mask the subtle genetic and epigenetic signatures of a nascent tumor [158,169].

Furthermore, the very technologies used for analysis, such as next-generation sequencing (NGS), introduce their own errors, which can lead to high false-positive rates if not properly managed [157]. The journey of a sample from patient to processor is also fraught with peril; preanalytical variables, including methods for collection, processing, and storage, can significantly impact the quality and integrity of delicate nucleic acid molecules, leading to analyte degradation and unreliable results [158,163,171]. Saliva, while advantageous owing to its proximity to oral tumors, presents unique challenges, as it contains a complex mixture of exfoliated cells, microorganisms, and enzymes that can interfere with analysis [162].

To transform SalivaSense 2.0 into a clinically robust tool, these pitfalls must be meticulously addressed. The field is rapidly moving toward ultrasensitive detection methods, such as digital PCR (dPCR) and digital sequencing, which employ unique molecular identifiers (UMIs) to tag each original DNA/RNA molecule [156,169]. This innovation allows for computational correction of amplification and sequencing errors, enabling the accurate quantification of true, low-frequency variants against background noise [156,157]. However, technology alone is not a panacea. The development of rigorously standardized and validated protocols for every preanalytical step, from the design of the saliva collection kit to the temperature of the freezer, is paramount to ensuring the reproducibility and reliability of results across different laboratories and patient populations [158,171]. Overcoming these analytical hurdles is the critical step that will build the foundation of the trust required for SalivaSense 2.0's widespread adoption in clinical care.

7.2. Data Privacy, Genomic Sovereignty and Health-Equity Safeguards

The successful implementation of SalivaSense 2.0 will generate vast repositories of deeply personal genomic data, creating a profound responsibility to protect individuals and communities while advancing science [165,172]. The history of genomics is, unfortunately, marked by instances of "biopiracy" and "biocoloniality," where genetic resources from indigenous peoples and populations in developing nations were collected and used by foreign entities without proper consent or fair sharing of the resulting benefits [165,166]. This legacy of exploitation has given rise to the concept of genomic sovereignty, a framework through which nations and indigenous communities assert their right to govern data derived from their people and lands [165,166,172].

Policies rooted in genomic sovereignty, such as those enacted in China and Mexico, aim to prevent uncompensated exploitation by placing strict controls on the export and use of "human genetic resources" [165,172]. While these measures are necessary responses to historical injustices, they present a modern paradox. In an era where biomedical breakthroughs depend on analyzing massive and diverse datasets, overly restrictive data-sharing policies can create information silos [172]. This risks worsening the existing underrepresentation of non-European populations in global genomic databases, which in turn limits the applicability of new diagnostics and therapies for these very communities, thereby undermining the goal of health equity [165].

Parallel to these collective rights are the fundamental individual rights to data privacy and security [172]. The European Union's General Data Protection Regulation (GDPR) exemplifies a stringent, rights-based approach, establishing strict rules for processing and transferring sensitive genetic data that create significant hurdles for international research collaborations [172]. These intersecting frameworks of national interest and individual rights highlight a central tension in global genomics. To move forward, a paradigm shift is needed, from a possessive model of genomic sovereignty to a collaborative ethos of genomic solidarity [165]. This approach reframes genetic differences not as a basis for exclusion but as a catalyst for inclusive and reciprocal partnerships. It calls for robust governance systems that ensure that data are shared responsibly, with transparent consent processes and mechanisms that guarantee the benefits of research, whether clinical, educational, or economic, and flow back to the communities that contribute their data [164,166].

7.3. Global Consortia, Capacity Building and Open-Science Models

Navigating the intricate technical and ethical landscapes of genomic medicine requires a united global effort. The vision of a "genomic commons", a worldwide, interoperable ecosystem of genomic and health data governed by shared principles of access and ethics, offers a powerful model for accelerating discovery [172]. However, this commons does not arise spontaneously; it must be actively and deliberately constructed through international collaboration, shared standards, and a commitment to open science [164].

Pioneering organizations such as the Global Alliance for Genomics and Health (GA4GH) and the International Liquid Biopsy Standardization Alliance (ILSA) are at the forefront of this construction effort [164,171]. They bring together researchers, clinicians, patient advocates, and industry partners to forge the tools necessary for responsible global data sharing. This work involves creating technical standards, such as common file formats, data models, and application programming interfaces (APIs), that ensure that different datasets can "speak" to one another, as well as developing policy frameworks and tools for ethics review, data access, and consent that are respected across borders [164,171]. A key innovation is the promotion of federated data analysis, a model in which researchers send their analytical tools to the data at a secure location rather than moving large, sensitive datasets across jurisdictions. This approach respects data privacy and sovereignty while still enabling the powerful, large-scale analyses needed to make discoveries [164].

However, technical and ethical frameworks alone are insufficient to guarantee global equity. True progress requires a profound investment in research capacity building in historically underresourced regions [167,172]. Initiatives such as the Diversity Centers for Genome Research (DCGR) Consortium exemplify this commitment by focusing on strengthening skills, developing infrastructure, and fostering partnerships at institutions that serve underrepresented communities [167]. Such efforts ensure that all nations can participate as equal partners in the genomic revolution, not merely as sources of data. This collaborative, open-science model, founded on standardized tools, shared ethical principles, and a dedication to building capacity worldwide, provides the essential foundation for translating powerful technologies such as SalivaSense 2.0 into tangible, equitable, and sustainable improvements in global human health.

8. Future Horizons and Innovation Roadmap

The successful detection of cfDNA and cfRNA from saliva marks a pivotal moment in oncology, yet it is merely the dawn of a new era in noninvasive diagnostics [173–193]. The true frontier lies not in refining a single modality but in weaving together a rich tapestry of biological data with the threads of advanced technology. The innovation roadmap for SalivaSense 2.0 and beyond envisions a fully integrated ecosystem, one that moves diagnostics from the centralized lab into the patient's home, transforms static data points into dynamic risk profiles, and leverages artificial intelligence to create a continuously learning, adaptive surveillance system. This future is built upon four interconnected pillars: (i) multimodal sensing, (ii) real-time edge-cloud analytics, (iii) patient-centric tele-oncology, and (iv) AI-driven discovery. This vision recognizes that genomics and transcriptomics, while powerful, represent only the visible 'tip of the iceberg' in a vast ocean of biological data. The true potential for a paradigm shift in diagnostics lies beneath the surface in the integrated analysis of multiple data layers through artificial intelligence, as conceptualized in (Figure 10) [116].

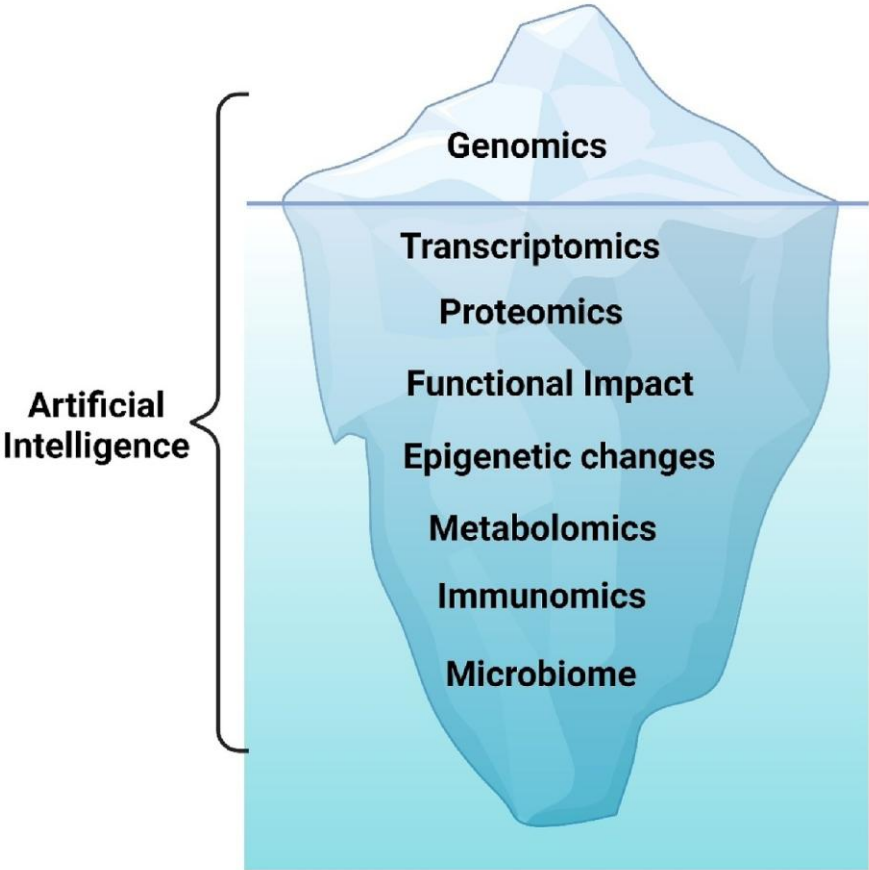


Figure 10. AI-integrated multimodal diagnostic landscape. The future of noninvasive diagnostics lies beyond genomics, which represents only the 'tip of the iceberg.' A comprehensive understanding of disease requires the integration of deeper layers of biological information, including the transcriptome, proteome, metabolome, and microbiome. Artificial intelligence is the key enabling technology that can analyze these complex, high-dimensional datasets to uncover hidden patterns and create a holistic molecular picture of health and disease. Reprinted from [116] under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

8.1. Multimodal Saliva Diagnostics: Microbiome, Volatilome, and Proteo-Glycomic Sensors

While cfDNA and cfRNA provide a direct line to the tumor's genomic and transcriptomic state, they tell only part of the story [186]. The oral cavity is a complex ecosystem where tumors influence ripples through every component of saliva [174,189]. The next evolution of SalivaSense will therefore be a multimodal platform, integrating cfDNA/cfRNA analysis with sensors that capture the broader biological context [193]. Research has revealed that the salivary microbiome of OSCC patients is profoundly altered, with a significant increase in diversity and enrichment of specific bacterial genera, such as *Streptococcus*, *Lactobacillus*, and *Prevotella* [189]. A future device could incorporate 16S rRNA sequencing or targeted molecular probes to detect these microbial signatures, adding a powerful layer of diagnostic information [191].

Simultaneously, the salivary proteome offers a window into the host's response to cancer [187]. Advanced proteomic analyses have identified dozens of differentially abundant proteins, including heat shock 70 kDa protein 1A/1B and key glycolytic enzymes, whose levels are significantly altered in OSCC patients [187]. These protein biomarkers, along with key cytokines such as IL-6, which is a more reliable indicator of OSCC progression in saliva than in serum, can be monitored via highly sensitive techniques [190,191]. The integration of advanced proteoglycomic sensors would allow SalivaSense 2.0 to capture not only the presence of the tumor's genetic material but also the functional impact it has on the host's immune and metabolic systems [187]. By combining these diverse data

streams, genomic, transcriptomic, microbial, and proteomic, we can construct a highly specific and robust molecular fingerprint of OSCC, dramatically improving early detection accuracy and our understanding of the intricate biology of this disease [175,189].

8.2. Real-Time Cloud-Edge Analytics and Risk Dashboards

The torrent of data generated by a multimodal SalivaSense 2.0 platform necessitates a paradigm shift in how we process and interpret health information. Sending raw, high-dimensional data to a centralized cloud for analysis is not only slow but also raises significant privacy concerns, especially when dealing with sensitive health records [181,182]. The future lies in a hybrid cloud-edge computing architecture, a distributed model that brings computation closer to the point of data generation [181]. In this model, an edge computing layer, residing on a local device or a nearby server, performs initial data processing [181,182]. This localized processing offers two transformative benefits: it drastically reduces latency, enabling real-time analysis critical for timely clinical decisions, and it enhances security and privacy by minimizing the transfer of sensitive raw data over the network [181].

Once processed at the edge, anonymized and aggregated insights can be securely transmitted to a central cloud platform [181]. This cloud layer serves as the engine for more computationally intensive tasks, such as training large-scale AI models on data from thousands of patients and identifying population-level trends [182]. The ultimate output for the clinician would be a dynamic, real-time risk dashboard. This intuitive interface can visualize a patient's integrated biomarker profile, track changes over time, and flag anomalies that may indicate disease progression or recurrence [177]. This system transforms diagnostics from a single-point-in-time event into a continuous monitoring process, empowering clinicians with actionable, data-driven insights for proactive and personalized patient care [181].

8.3. At-Home Smart Cup Sampling Linked to Tele-Oncology

To truly revolutionize OSCC screening and surveillance, diagnostics must escape the confines of the clinic and enter the patient's daily life. Owing to the noninvasive and simple nature of saliva collection, it is uniquely suited for at-home use, eliminating the pain, inconvenience, and risk associated with blood draws [174,175,180]. The next generation of SalivaSense will incorporate this principle through the development of a "smart-cup" or similar minimally instrumented, smartphone-based device for automated sample collection and stabilization [192]. This device is designed for ease of use, requires no specialized training and allows patients to provide samples from the comfort of their homes [192].

This at-home sampling capability is a perfect catalyst for integration with emerging tele-oncology platforms. Telehealth frameworks, which are becoming increasingly sophisticated, provide the infrastructure for remote patient monitoring and consultation [173]. A patient using a smart-cup device could have their sample data automatically and securely uploaded to a clinician's dashboard via their smartphone [192]. This link between at-home diagnostics and virtual care creates a powerful feedback loop, enabling continuous surveillance without the need for frequent and burdensome hospital visits [173]. For patients in rural or underserved areas, this model dramatically improves access to specialized oncological care [173]. The healthcare system represents a more efficient and scalable model for managing chronic disease surveillance, reducing costs while improving patient compliance and quality of life [174,192].

8.4. AI-Driven Discovery and Adaptive Surveillance Paradigms

The culmination of these advancements is an intelligent, self-improving healthcare ecosystem driven by artificial intelligence (AI). The role of AI in this future horizon extends far beyond simple binary classification of "cancer" or "no cancer." By applying advanced deep learning models to rich, multimodal datasets generated by SalivaSense 2.0, we can begin to uncover previously hidden

biological patterns and discover novel biomarkers for OSCC [183]. This AI-driven discovery engine functions as a continuous feedback loop; as more patient data are collected, the AI models become more adept at identifying subtle molecular signatures associated with early-stage disease, treatment resistance, or risk of recurrence [177,183]. Federated learning approaches would allow these models to be trained across multiple institutions without compromising patient privacy, accelerating the pace of discovery on a global scale [183].

This powerful analytical capability enables a truly adaptive surveillance paradigm. Instead of relying on rigid, one-size-fits-all monitoring schedules, AI algorithms can create personalized surveillance plans for each patient [176]. By continuously analyzing a patient's real-time biomarker data from their smart cup, the system can dynamically adjust the frequency of testing on the basis of their evolving risk profile [179]. A patient with a stable molecular signature might have their follow-up extended, whereas a patient with emerging signs of molecular recurrence would be flagged for earlier clinical intervention [184]. This AI-guided approach ensures that clinical resources are directed where they are most needed, maximizing efficiency and improving patient outcomes. The implementation of such a powerful system, however, must be governed by robust risk management frameworks to ensure that its use is safe, equitable, and aligned with core ethical principles [178,179].

9. Conclusion

Oral squamous cell carcinoma (OSCC) remains a formidable global health challenge, not because of a lack of effective treatments but because of persistent shortfalls in early diagnosis. The stark difference in five-year survival rates between localized and advanced-stage disease, which decreases from as high as 90% to less than 50%, highlights a grim picture of a crisis defined by delayed detection. This diagnostic void is most pronounced in the very low-resource and community settings that bear the heaviest burden of mortality, creating an urgent global mandate to shift the paradigm from reactive surgical intervention toward proactive, molecularly guided prevention. This review has articulated the case for SalivaSense 2.0, a conceptual next-generation framework designed to meet this challenge by transforming routine saliva sampling into an ultrasensitive, multiplexed liquid biopsy capable of detecting OSCC at its molecular dawn.

The foundational innovation of SalivaSense 2.0 lies in its move away from single-analyte biomarkers toward an integrated epigenomic approach. By synergistically combining the analysis of stable, tumor-specific cell-free DNA (cfDNA) methylation patterns with dynamic, real-time snapshots of cell-free RNA (cfRNA) transcriptional profiles, this strategy generates a comprehensive and robust fingerprint of pico-stage tumorigenesis. Realizing this vision requires a sophisticated and seamlessly integrated technological pipeline. At its core is an automated microfluidic cartridge engineered to overcome the profound challenge of isolating picogram-level nucleic acids from the complex salivary milieu with exceptional efficiency and purity. Downstream, this captured material is interrogated by a dual analytical engine: direct, long-read nanopore sequencing, which offers an unprecedented, conversion-free view of the native epigenome, and ultrasensitive electrochemical nanobiosensors, which pave the way for rapid, point-of-care quantification of key validated markers.

This technological architecture is powered and unified by a sophisticated artificial intelligence pipeline that serves as the framework's central nervous system. AI is not merely an accessory but a foundational pillar that is essential for distilling a clear diagnostic signal from immense biological and technical noise. This begins with preprocessing and error suppression, where techniques such as integrated digital error suppression (iDES) combine molecular barcoding with computational background polishing to eliminate sequencing artifacts, pushing the limits of detection to mutant allele frequencies as low as 4×10^{-5} . Subsequently, adaptive baseline algorithms have been deployed to mitigate the effects of biological confounders, from circadian rhythms and tobacco metabolites to the molecular noise generated by the oral microbiome and periodontal inflammation, thus preserving the integrity of the pico-scale cancer signal. Finally, advanced machine learning models, such as graph neural networks and pathway-aware transformers, move beyond simplistic analysis to perform intelligent multiomic integration. These models learn the complex, synergistic relationships

between cfDNA and cfRNA features, dramatically enhancing predictive accuracy while providing invaluable, explainable insights into the biological pathways driving the disease.

The journey from a validated algorithm to a trusted clinical tool is paved with rigorous validation and equitable implementation. The SalivaSense 2.0 framework necessitates a multistage validation pathway, beginning with analytical benchmarking against certified reference standards under dMIQE guidelines and proceeding to blinded, multicenter clinical trials that prove its performance against the gold standard of histopathology. Crucially, health-economic modeling and decision curve analysis are required to demonstrate not only accuracy but also tangible clinical utility and cost effectiveness, ensuring that the technology provides clear value to patients and healthcare systems. The ultimate goal is to translate this power into a truly accessible point-of-care diagnostic. This demands engineering innovations such as smartphone-linked optical readers and electricity-free heating modules, coupled with logistical solutions such as dried-saliva-spot transport that democratize access for remote and underserved populations. This must be accompanied by a human-centered implementation strategy that includes workforce training, culturally resonant community engagement, and robust data governance pathways, such as federated learning and dynamic consent, that honor both patient privacy and global genomic sovereignty. By integrating these technological, clinical, and social imperatives, a framework such as SalivaSense 2.0 can fulfill its promise, catalyzing the decisive shift from managing advanced OSCC to proactively preventing it, thereby dramatically improving patient outcomes worldwide.

10. Unresolved Questions and Future Research Directives

While the roadmap for developing and implementing advanced salivary diagnostics is becoming clearer, the frontier of this field is defined by a series of profound and unresolved questions. Answering these questions will not only refine existing technologies but also unlock entirely new paradigms in noninvasive disease detection.

First, we must investigate the fundamental biology of biomarker release at the earliest stages of carcinogenesis. Beyond simply measuring tumor burden, what are the intrinsic molecular and microenvironmental factors that determine why some premalignant lesions are "loud shedders" of cfNAs while others remain molecularly silent? Does a "shedding phenotype", governed by lesion vascularity, apoptotic flux, or specific oncogenic driver pathways, exist and precede clinically detectable cancer? The development of methods to quantify this shedding efficiency itself could yield a novel, orthogonal risk stratification tool that acts as a sentinel for malignant transformation, independent of the specific mutation being sought.

Second, the current approach treats biological confounders such as periodontal inflammation and microbial dysbiosis primarily as noise to be corrected. The next step is to reframe this "noise" as a rich source of contextual information. Can AI models be trained not only to subtract signals from gingival crevicular fluid but also to recognize specific host-microbiome interaction patterns that are uniquely permissive of carcinogenesis? The critical question is whether the epigenetic signature of inflammation in a patient with high-risk oral lesions is measurably different from that in a low-risk individual with the same periodontal condition. If so, we could transform a major confounder into a context-dependent biosensor, harnessing its signal to dramatically improve diagnostic specificity.

Finally, as deep learning models become more powerful, we face the challenge of reverse translation. When a model identifies a highly predictive but biologically nonintuitive feature, a "dark biomarker" from the AI's "black box", what is the optimal pipeline to establish its functional role? We must move beyond explainability dashboards and build frameworks for "AI-guided biology." Can the internal attention maps of a transformer model be used to directly design CRISPR screens or functional assays in organoid models? This would create a virtuous cycle where the AI not only provides a diagnosis but also actively directs biological research toward novel oncogenic mechanisms, fundamentally accelerating the pace of discovery.

List of Abbreviations

Abbreviation	Full Term
5hmC	5-Hydroxymethylcytosine
5mC	5-Methylcytosine
AI	Artificial Intelligence
API	Application Programming Interface
AR	Augmented Reality
AutoML	Automated Machine Learning
AuNP	Gold Nanoparticle
bp	Base Pair
ccfDNA	Circulating Cell-Free DNA
cfDNA	Cell-Free DNA
cfNA	Cell-Free Nucleic Acid
cfRNA	Cell-Free RNA
ConQuR	Conditional Quantile Regression
CRM	Certified Reference Material
CTC	Circulating Tumor Cell
ctDNA	Circulating Tumor DNA
ctNA	Circulating Tumor Nucleic Acid
ctRNA	Circulating Tumor RNA
DALY	Disability-Adjusted Life Year
DCA	Decision Curve Analysis
DCP	Dynamic Consent Platform
DCGR	Diversity Centers for Genome Research
ddPCR	Droplet Digital PCR
dMIQE	Minimum Information for Publication of Quantitative Digital PCR Experiments
dncfDNA	Dinucleosomal Cell-Free DNA
DSS	Dried Saliva Spot
ELISA	Enzyme-Linked Immunosorbent Assay
EU	European Union
EV	Extracellular Vesicle
FDA	Food and Drug Administration
FL	Federated Learning
FTL	Federated Transfer Learning
GA4GH	Global Alliance for Genomics and Health
GAIN	Generative Adversarial Imputation Network
GAT	Graph Attention Network
GCF	Gingival Crevicular Fluid
GCN	Graph Convolutional Network
GDPR	General Data Protection Regulation
GNN	Graph Neural Network
ICER	Incremental Cost-Effectiveness Ratio
iDES	Integrated Digital Error Suppression
IL-1β	Interleukin-1 Beta
ILSA	International Liquid Biopsy Standardization Alliance
IVDR	In Vitro Diagnostic Medical Devices Regulation
kb	Kilobase

LMIC	Low- and Middle-Income Country
LOD	Limit of Detection
miRNA	microRNA
mncfDNA	Mononucleosomal Cell-Free DNA
mRNA	Messenger RNA
NGS	Next-Generation Sequencing
OPMD	Oral Potentially Malignant Disorder
OSCC	Oral Squamous Cell Carcinoma
oxBS-seq	Oxidative Bisulfite Sequencing
PCM	Phase Change Material
PGV	Pathogenic Germline Variant
POC	Point-of-Care
PPV	Positive Predictive Value
PQDx	Prequalification of In Vitro Diagnostics
QALY	Quality-Adjusted Life Year
RCA	Rolling-Circle Amplification
rGO	Reduced Graphene Oxide
RFS	Robust Feature Selection
ROS	Reactive Oxygen Species
SCCA	Squamous Cell Carcinoma Antigen
scfDNA	Salivary Cell-Free DNA
scRNA-seq	Single-Cell RNA Sequencing
SDA	Strand Displacement Amplification
SDI	Sociodemographic Index
SI	International System of Units
SOP	Standardized Operating Procedure
SPE	Solid-Phase Extraction
SPDF	Secure Product Development Framework
TAB-seq	TET-Assisted Bisulfite Sequencing
tncfDNA	Trinucleosomal Cell-Free DNA
UMI	Unique Molecular Identifier
uscfDNA	Ultrashort Cell-Free DNA
WHO	World Health Organization

Author information: Abdullah Ayad: A distinguished molecular biologist and geneticist at the forefront of research and development in molecular diagnostics and genetics. Currently serving as a staff scientist at the Al Bayan National Laboratory for advanced medical diagnostics, Abdullah's work spans several critical areas, including molecular biology, genetics, oncology, and biotechnology. His expertise is particularly notable in the fields of next-generation sequencing (NGS) and CRISPR/Cas9 gene editing technologies. Abdullah's academic journey in dentistry and his subsequent pivot to molecular genetics illustrate his commitment to integrating diverse fields of science to address complex health issues. He is actively involved in pioneering research on genetic diseases, such as cancer and Sandhoff syndrome, utilizing cutting-edge technologies to explore novel therapeutic strategies and diagnostics. With a strong focus on collaboration and knowledge dissemination, Abdullah has established significant partnerships both locally and internationally, engaging in collaborative projects and conferences that highlight his commitment to advancing the field of molecular genetics. His work with the Ministry of Health in Iraq and his efforts to introduce new technologies such as the Oxford Nanopore Technologies' MinION devices are testament to his role as a key player in expanding the capabilities of genetic research and diagnostics in the region. As an academic and researcher, Abdullah has contributed to numerous scholarly articles and has been an active participant in peer reviews, further enriching the scientific community. His vision for the future includes leading a diagnostics and R&D center that not only focuses on research but also serves as a hub for innovation in genetic diagnostics and therapy in Iraq. **Yousef Husam:** A rising Iraqi

dental researcher whose work bridges clinical dentistry and molecular genetics. Based at the College of Dentistry, University of Basrah, he is a core member of the Basra Scientific Research Team Lab, coordinating projects that examine how genetic variants influence oral development, pathology, and tissue regeneration. Since 2023 he has co-authored seven peer-reviewed articles and preprints that have attracted almost two thousand reads, contributing to advances in areas such as radiation safety in dental imaging, nano-enabled tissue regeneration, and the molecular basis of inherited oral disorders. His scholarly work synthesizes advances in high-resolution imaging, biomimetic scaffolds, and nanotherapeutics to chart a roadmap for next-generation dental care. Beyond scholarship, Hussam mentors undergraduate students in research methodology and liaises with local clinics to pilot precision-dentistry workflows

- Funding: Not Applicable
- Ethics approval and consent to participate: Not Applicable
- Consent for publication: Not Applicable
- Availability of data and material: Not Applicable
- Acknowledgments: Not Applicable
- Competing interests: Not Applicable

Declaration Regarding the Use of AI-Assisted Readability Enhancement: We hereby affirm that AI tools were used only for initial drafting to improve readability, grammar, language quality, and document organization. At no point did they replace authorial responsibilities or independent judgments. AI involvement was rigorously supervised under the discerning eye of human oversight and control. AI works within a closed corpus limited to the peer-reviewed sources cited, summarizing passages, proposing outlines, and suggesting draft phrasing with source tags. These suggestions served only as prompts. The authors reviewed every line, verified the text and citations, ensured that each retained statement mapped to a specific citation, and approved the final wording, analyses, and conclusions. No generative AI images or videos were used.

References

1. Wang, Q., Gao, P., Wang, X., & Duan, Y. (2014). Investigation and identification of potential biomarkers in human saliva for the early diagnosis of oral squamous cell carcinoma. *Clinica Chimica Acta*, 427, 79-85.
2. Saleh, A., Kong, Y. H., Haron, N., Aripin, S. F., Vadiveloo, M., Hussaini, H., ... & Cheong, S. C. (2017). Oral cancer screening in private dental practices in a developing country: opportunities and challenges. *Community dentistry and oral epidemiology*, 45(2), 112-119.
3. Hu, C. C., Wang, S. G., Gao, Z., Qing, M. F., Pan, S., Liu, Y. Y., & Li, F. (2025). Emerging salivary biomarkers for early detection of oral squamous cell carcinoma. *World Journal of Clinical Oncology*, 16(4), 103803.
4. Semprini, J., Reynolds, J., Zahnd, W. E., & Wehby, G. (2024). Eliminating Medicaid dental benefits and early-stage oral cancer diagnoses. *Cancer Medicine*, 13(4), e7061.
5. Chen, M., Li, J., Su, W., Huang, J., Yang, C., Li, R., & Chen, G. (2025). Global Burden of Lip and Oral Cavity Cancer From 1990 to 2021 and Projection to 2040: Findings From the 2021 Global Burden of Disease Study. *Cancer Medicine*, 14(9), e70957.
6. Liu, L., Zhong, X., Zhong, Y., & Li, L. (2025). Recent advances in biomarker detection of oral squamous cell carcinoma. *Frontiers in Oncology*, 15, 1597086.
7. Patel, A., Patel, S., Patel, P., & Tanavde, V. (2022). Saliva based liquid biopsies in head and neck cancer: how far are we from the clinic?. *Frontiers in oncology*, 12, 828434.
8. Cristaldi, M., Mauceri, R., Di Fede, O., Giuliana, G., Campisi, G., & Panzarella, V. (2019). Salivary biomarkers for oral squamous cell carcinoma diagnosis and follow-up: current status and perspectives. *Frontiers in physiology*, 10, 1476.
9. Saidak, Z., Milly, M., Louandre, C., Colin, E., Rusu, P. M., Paasche, A., ... & Galmiche, A. (2025). Salivary cell-free DNA methylation analysis for oncological monitoring of surgical resection of oral squamous cell carcinoma. *Frontiers in Oral Health*, 6, 1614371.
10. Le Campion, A. C. O. V., Ribeiro, C. M. B., Luiz, R. R., da Silva Júnior, F. F., Barros, H. C. S., Dos Santos, K. D. C. B., ... & Ferreira, S. M. S. (2017). Low survival rates of oral and oropharyngeal squamous cell carcinoma. *International journal of dentistry*, 2017(1), 5815493.

11. Al-Maweri, S. A., Halboub, E., Shamala, A., Al-Maweri, A. A., Daud, A., Almurisi, E., ... & Alhajj, M. N. (2025). Oral Cancer Knowledge and Screening Practices Among Dental Professionals in Yemen: a Web-Based Survey. *International Dental Journal*, 75(3), 2034-2041.
12. Nassani, M. Z., Alsalthani, A., Alali, F. M., Rastam, S., Alqhtani, N. R., Alqahtahni, A. S., ... & Tarakji, B. (2025). Public awareness and knowledge of oral cancer in 13 Middle Eastern and North African Countries. *JAMA Network Open*, 8(3), e250522-e250522.
13. Yang, E. C., Schwarz, R. A., Lang, A. K., Bass, N., Badaoui, H., Vohra, I. S., ... & Richards-Kortum, R. R. (2018). In vivo multimodal optical imaging: improved detection of oral dysplasia in low-risk oral mucosal lesions. *Cancer Prevention Research*, 11(8), 465-476.
14. Ribeiro-Rotta, R. F., Rosa, E. A., Milani, V., Dias, N. R., Masterson, D., da Silva, E. N., & Zara, A. L. D. S. A. (2022). The cost of oral cancer: A systematic review. *PLoS One*, 17(4), e0266346.
15. Hu, Y., Xu, M., Liu, M., & Peng, H. (2025). Comparison of saliva and blood derived cell free RNAs for detecting oral squamous cell carcinoma. *Scientific Reports*, 15(1), 4645.
16. Ghanem, A. S., Tóth, Á., & Nagy, A. C. (2025). Sociodemographic disparities in global trends of lip and oral cavity neoplasms from 1990 to 2021. *Scientific Reports*, 15(1), 4230.
17. Swarup, N., Cheng, J., Choi, I., Heo, Y. J., Kordi, M., Aziz, M., ... & Wong, D. T. (2023). Multifaceted attributes of salivary cell-free DNA as liquid biopsy biomarkers for gastric cancer detection. *Biomarker research*, 11(1), 90.
18. Cui, L., Zheng, J., Lu, Y., Lin, P., Lin, Y., Zheng, Y., ... & Zhao, X. (2024). New frontiers in salivary extracellular vesicles: transforming diagnostics, monitoring, and therapeutics in oral and systemic diseases. *Journal of Nanobiotechnology*, 22(1), 171.
19. Swarup, N., Leung, H. Y., Choi, I., Aziz, M. A., Cheng, J. C., & Wong, D. T. (2025). Cell-Free DNA: Features and Attributes Shaping the Next Frontier in Liquid Biopsy. *Molecular Diagnosis & Therapy*, 29(3), 277-290.
20. Rigotti, P., Polizzi, A., Quinzi, V., Blasi, A., Lombardi, T., Lo Muzio, E., & Isola, G. (2025). Cell-Free DNA as a Prognostic Biomarker in Oral Carcinogenesis and Oral Squamous Cell Carcinoma: A Translational Perspective. *Cancers*, 17(14), 2366.
21. Park, N. J., Li, Y., Yu, T., Brinkman, B. M., & Wong, D. T. (2006). Characterization of RNA in saliva. *Clinical chemistry*, 52(6), 988-994.
22. Bronkhorst, A. J., Ungerer, V., Oberhofer, A., Gabriel, S., Polatoglou, E., Randeu, H., ... & Holdenrieder, S. (2022). New perspectives on the importance of cell-free DNA biology. *Diagnostics*, 12(9), 2147.
23. Salfer, B., Havo, D., Kuppinger, S., Wong, D. T., Li, F., & Zhang, L. (2023). Evaluating preanalytical variables for saliva cell-free DNA liquid biopsy. *Diagnostics*, 13(10), 1665.
24. Janovičová, L., Holániová, D., Vlková, B., & Celec, P. (2024). Pre-Analytical Factors Affecting Extracellular DNA in Saliva. *Diagnostics*, 14(3), 249.
25. De, S. (2021). Signatures beyond oncogenic mutations in cell-free DNA sequencing for noninvasive, early detection of cancer. *Frontiers in Genetics*, 12, 759832.
26. Saidak, Z., Milly, M., Louandre, C., Colin, E., Rusu, P. M., Paasche, A., ... & Galmiche, A. (2025). Salivary cell-free DNA methylation analysis for oncological monitoring of surgical resection of oral squamous cell carcinoma. *Frontiers in Oral Health*, 6, 1614371.
27. Li, J. J., Liu, G., & Lok, B. H. (2024). Cell-Free DNA Hydroxymethylation in Cancer: Current and Emerging Detection Methods and Clinical Applications. *Genes*, 15(9), 1160.
28. Szilágyi, M., Pös, O., Márton, É., Buglyó, G., Soltész, B., Keserű, J., ... & Nagy, B. (2020). Circulating cell-free nucleic acids: main characteristics and clinical application. *International journal of molecular sciences*, 21(18), 6827.
29. Pös, Z., Pös, O., Styk, J., Mocova, A., Strieskova, L., Budis, J., ... & Szemes, T. (2020). Technical and methodological aspects of cell-free nucleic acids analyzes. *International Journal of Molecular Sciences*, 21(22), 8634.
30. Hyun, K. A., Gwak, H., Lee, J., Kwak, B., & Jung, H. I. (2018). Salivary exosome and cell-free DNA for cancer detection. *Micromachines*, 9(7), 340.
31. Zhang, X. (2025). Highly effective batch effect correction method for RNA-seq count data. *Computational and Structural Biotechnology Journal*, 27, 58-64.

32. Kotlov, N., Shaposhnikov, K., Tazearslan, C., Chasse, M., Baisangurov, A., Podsvirova, S., ... & Bagaev, A. (2024). Procrustes is a machine-learning approach that removes cross-platform batch effects from clinical RNA sequencing data. *Communications Biology*, 7(1), 392.
33. Bronkhorst, A. J., Ungerer, V., Oberhofer, A., Gabriel, S., Polatoglou, E., Randeu, H., ... & Holdenrieder, S. (2022). New perspectives on the importance of cell-free DNA biology. *Diagnostics*, 12(9), 2147.
34. Salfer, B., Havo, D., Kuppinger, S., Wong, D. T., Li, F., & Zhang, L. (2023). Evaluating preanalytical variables for saliva cell-free DNA liquid biopsy. *Diagnostics*, 13(10), 1665.
35. Galvin, S., Moran, G. P., & Healy, C. M. (2023). Influence of site and smoking on malignant transformation in the oral cavity: Is the microbiome the missing link?. *Frontiers in Oral Health*, 4, 1166037.
36. Kobayashi, H., Song, C., Ikei, H., Park, B. J., Kagawa, T., & Miyazaki, Y. (2017). Diurnal changes in distribution characteristics of salivary cortisol and immunoglobulin A concentrations. *International journal of environmental research and public health*, 14(9), 987.
37. Reis, M. G., Lopes, L. C., Sanches, A. B. A. M. D. A., Guimarães, N. S., & Martins-Chaves, R. R. (2024). Diet and Oral Squamous Cell Carcinoma: A Scoping Review. *International Journal of Environmental Research and Public Health*, 21(9), 1199.
38. Oyeyemi, B. F., Kaur, U. S., Paramraj, A., Tandon, R., Kumar, A., & Bhavesh, N. S. (2023). Microbiome analysis of saliva from oral squamous cell carcinoma (OSCC) patients and tobacco abusers with potential biomarkers for oral cancer screening. *Heliyon*, 9(11).
39. Rajasekaran, J. J., Krishnamurthy, H. K., Bosco, J., Jayaraman, V., Krishna, K., Wang, T., & Bei, K. (2024). Oral microbiome: a review of its impact on oral and systemic health. *Microorganisms*, 12(9), 1797.
40. Chellappa, S. L., Engen, P. A., Naqib, A., Qian, J., Vujovic, N., Rahman, N., ... & Scheer, F. A. (2022). Proof-of-principle demonstration of endogenous circadian system and circadian misalignment effects on human oral microbiota. *The FASEB Journal*, 36(1), e22043.
41. Wang, S., Tan, X., Cheng, J., Liu, Z., Zhou, H., Liao, J., ... & Liu, H. (2024). Oral microbiome and its relationship with oral cancer. *Journal of Cancer Research and Therapeutics*, 20(4), 1141-1149.
42. Sprang, M., Andrade-Navarro, M. A., & Fontaine, J. F. (2022). Batch effect detection and correction in RNA-seq data using machine-learning-based automated assessment of quality. *BMC bioinformatics*, 23(Suppl 6), 279.
43. Behdenna, A., Colange, M., Haziza, J., Gema, A., Appé, G., Azencott, C. A., & Nordor, A. (2023). pyComBat, a Python tool for batch effects correction in high-throughput molecular data using empirical Bayes methods. *BMC bioinformatics*, 24(1), 459.
44. Yu, P., Chen, P., Wu, M., Ding, G., Bao, H., Du, Y., ... & Cheng, X. (2024). Multidimensional cell-free DNA-based liquid biopsy for sensitive early detection of gastric cancer. *Genome Medicine*, 16(1), 79.
45. Ling, W., Lu, J., Zhao, N., Lulla, A., Plantinga, A. M., Fu, W., ... & Wu, M. C. (2022). Batch effects removal for microbiome data via conditional quantile regression. *Nature communications*, 13(1), 5418.
46. Sarkar, A., Kuehl, M. N., Alman, A. C., & Burkhardt, B. R. (2021). Linking the oral microbiome and salivary cytokine abundance to circadian oscillations. *Scientific reports*, 11(1), 2658.
47. Poulet, G., Hulot, J. S., Blanchard, A., Bergerot, D., Xiao, W., Ginot, F., ... & Taly, V. (2023). Circadian rhythm and circulating cell-free DNA release on healthy subjects. *Scientific Reports*, 13(1), 21675.
48. Nath, S., Zilm, P., Jamieson, L., Santiago, P. H. R., Ketagoda, D. H. K., & Weyrich, L. (2025). The influence of diet, saliva, and dental history on the oral microbiome in healthy, caries-free Australian adults. *Scientific reports*, 15(1), 18755.
49. Nelson, N., Malhan, D., Hesse, J., Aboumanify, O., Yalçın, M., Lüers, G., & Relógio, A. (2025). Comprehensive integrative analysis of circadian rhythms in human saliva. *npj Biological Timing and Sleep*, 2(1), 17.
50. Moran, S. P., Rosier, B. T., Henriquez, F. L., & Burleigh, M. C. (2024). The effects of nitrate on the oral microbiome: a systematic review investigating prebiotic potential. *Journal of Oral Microbiology*, 16(1), 2322228.
51. Wang, Z., Chen, Y., Li, H., Yue, Y., & Yu, H. (2025). Exploring oral microbiome in oral squamous cell carcinoma across environment-associated sample types. *Microbiology Spectrum*, 13(4), e00852-24.

52. Obino, D., Vassalli, M., Franceschi, A., Alessandrini, A., Facci, P., & Viti, F. (2021). An overview on microfluidic systems for nucleic acids extraction from human raw samples. *Sensors*, 21(9), 3058.

53. Kumar, P., Gupta, S., & Das, B. C. (2024). Saliva as a potential noninvasive liquid biopsy for early and easy diagnosis/prognosis of head and neck cancer. *Translational oncology*, 40, 101827.

54. Lee, H., Park, C., Na, W., Park, K. H., & Shin, S. (2020). Precision cell-free DNA extraction for liquid biopsy by integrated microfluidics. *NPJ precision oncology*, 4(1), 3.

55. Brooks, P. J., Malkin, E. Z., De Michino, S., & Bratman, S. V. (2023). Isolation of salivary cell-free DNA for cancer detection. *Plos one*, 18(5), e0285214.

56. van Dessel, L. F., Vitale, S. R., Helmijr, J. C., Wilting, S. M., van der Vlugt-Daane, M., Oomen-de Hoop, E., ... & Lolkema, M. P. (2019). High-throughput isolation of circulating tumor DNA: a comparison of automated platforms. *Molecular oncology*, 13(2), 392-402.

57. Clark, T. A., Chung, J. H., Kennedy, M., Hughes, J. D., Chennagiri, N., Lieber, D. S., ... & Lipson, D. (2018). Analytical validation of a hybrid capture-based next-generation sequencing clinical assay for genomic profiling of cell-free circulating tumor DNA. *The Journal of Molecular Diagnostics*, 20(5), 686-702.

58. Babji, D., Nayak, R., Bhat, K., & Kotrashetti, V. (2019). Cell-free tumor DNA: emerging reality in oral squamous cell carcinoma. *Journal of Oral and Maxillofacial Pathology*, 23(2), 273-279.

59. Sathyanarayana, S. H., Spracklin, S. B., Deharvengt, S. J., Green, D. C., Instasi, M. D., Gallagher, T. L., ... & Tsongalis, G. J. (2025). Standardized Workflow and Analytical Validation of Cell-Free DNA Extraction for Liquid Biopsy Using a Magnetic Bead-Based Cartridge System. *Cells*, 14(14), 1062.

60. Huang, J., & Wang, L. (2019). Cell-free DNA methylation profiling analysis—technologies and bioinformatics. *Cancers*, 11(11), 1741.

61. Seong, H., Park, J., Bae, M., & Shin, S. (2022). Rapid and efficient extraction of cell-free DNA using homobifunctional crosslinkers. *Biomedicines*, 10(8), 1883.

62. Shanmugam, A., Hariharan, A. K., Hasina, R., Nair, J. R., Katragadda, S., Irusappan, S., ... & Izumchenko, E. (2021). Ultrasensitive detection of tumor-specific mutations in saliva of patients with oral cavity squamous cell carcinoma. *Cancer*, 127(10), 1576-1589.

63. Huang, Z., Dou, Y., Su, J., Li, T., & Song, S. (2023). Electrochemical biosensing methods for detecting epigenetic Modifications. *Chemosensors*, 11(8), 424.

64. Ma, Z., Hu, Y., Wang, L., Li, M., Li, C., Li, L., ... & Zheng, J. (2024). Recent advances in biosensors for analysis of DNA/RNA methylation. *Sensing and Bio-Sensing Research*, 46, 100690.

65. Yuen, Z. W. S., Srivastava, A., Daniel, R., McNevin, D., Jack, C., & Eyra, E. (2021). Systematic benchmarking of tools for CpG methylation detection from nanopore sequencing. *Nature communications*, 12(1), 3438.

66. Liu, Y., Rosikiewicz, W., Pan, Z., Jillette, N., Wang, P., Taghbalout, A., ... & Li, S. (2021). DNA methylation-calling tools for Oxford Nanopore sequencing: a survey and human epigenome-wide evaluation. *Genome biology*, 22(1), 295.

67. Lau, B. T., Almeda, A., Schauer, M., McNamara, M., Bai, X., Meng, Q., ... & Ji, H. P. (2023). Single-molecule methylation profiles of cell-free DNA in cancer with nanopore sequencing. *Genome Medicine*, 15(1), 33.

68. Song, D., Zhang, Z., Zheng, J., Zhang, W., & Cai, J. (2025). 5-Hydroxymethylcytosine modifications in circulating cell-free DNA: frontiers of cancer detection, monitoring, and prognostic evaluation. *Biomarker Research*, 13(1), 39.

69. Li, J. J., Liu, G., & Lok, B. H. (2024). Cell-Free DNA Hydroxymethylation in Cancer: Current and Emerging Detection Methods and Clinical Applications. *Genes*, 15(9), 1160.

70. Rapado-González, Ó., Salta, S., López-López, R., Henrique, R., Suárez-Cunqueiro, M. M., & Jerónimo, C. (2024). DNA methylation markers for oral cancer detection in nonand minimally invasive samples: a systematic review. *Clinical Epigenetics*, 16(1), 105.

71. van Ginkel, J. H., Huibers, M. M., van Es, R. J., de Bree, R., & Willems, S. M. (2017). Droplet digital PCR for detection and quantification of circulating tumor DNA in plasma of head and neck cancer patients. *BMC cancer*, 17(1), 428.

72. Wang, L., Shen, X., Wang, T., Chen, P., Qi, N., Yin, B., & Ye, B. (2020). A lateral flow strip combined with Cas9 nickase-triggered amplification reaction for dual food-borne pathogen detection. *Biosensors and Bioelectronics*, 165, 112364.
73. Zhou, W., Hu, L., Ying, L., Zhao, Z., Chu, P. K., & Yu, X. F. (2018). A CRISPR–Cas9-triggered strand displacement amplification method for ultrasensitive DNA detection. *Nature communications*, 9(1), 5012.
74. Kang, X., Wu, C., Alibakhshi, M. A., Liu, X., Yu, L., Walt, D. R., & Wanunu, M. (2023). Nanopore-based fingerprint immunoassay based on rolling circle amplification and DNA fragmentation. *ACS nano*, 17(6), 5412-5420.
75. Marcozzi, A., Jager, M., Elferink, M., Straver, R., van Ginkel, J. H., Peltenburg, B., ... & de Ridder, J. (2021). Accurate detection of circulating tumor DNA using nanopore consensus sequencing. *NPJ genomic medicine*, 6(1), 106.
76. Bao, J., Ding, K., & Zhu, Y. (2023). An electrochemical biosensor for detecting DNA methylation based on AuNPs/rGO/g-C3N4 nanocomposite. *Analytical biochemistry*, 673, 115180.
77. Zou, A., Zhu, X., Fu, R., Wang, Z., Wang, Y., Ruan, Z., Xianyu, Y., & Zhang, J. (2025). Harnessing nanomaterials for Next-Generation DNA methylation biosensors. *Small*.
78. Chen, X., Huang, J., Zhang, S., Mo, F., Su, S., Li, Y., ... & Zheng, J. (2019). Electrochemical biosensor for DNA methylation detection through hybridization chain-amplified reaction coupled with a tetrahedral DNA nanostructure. *ACS applied materials & interfaces*, 11(4), 3745-3752.
79. Kishore, R., Kille, S. N., Kishore, R., Chakraborty, G., Rajendra, V. K. J., D'Cruz, A., & Chakraborty, A. (2025). Noninvasive saliva-based profiling of TP53 codon 72 polymorphism in oral cancer patients by droplet digital PCR. *Gene Reports*, 102256.
80. Valous, N. A., Popp, F., Zörnig, I., Jäger, D., & Charoentong, P. (2024). Graph machine learning for integrated multiomics analysis. *British journal of cancer*, 131(2), 205-211.
81. Zhang, X., Chen, J., Wang, Y., Wang, X., Hu, J., Peng, J., ... & Wang, T. (2025). cfMethylPre: deep transfer learning enhances cancer detection based on circulating cell-free DNA methylation profiling. *Briefings in Bioinformatics*, 26(3), bbaf303.
82. LeDell, E., & Poirier, S. (2020, July). H2o automl: Scalable automatic machine learning. In *Proceedings of the AutoML Workshop at ICML (Vol. 2020, p. 24)*.
83. Arango-Argoty, G., Kipkogei, E., Stewart, R., Sun, G. J., Patra, A., Kagiampakis, I., & Jacob, E. (2025). Pretrained transformers applied to clinical studies improve predictions of treatment efficacy and associated biomarkers. *Nature Communications*, 16(1), 2101.
84. Liu, X., Shi, J., Jiao, Y., An, J., Tian, J., Yang, Y., & Zhuo, L. (2024). Integrated multiomics with machine learning to uncover the intricacies of kidney disease. *Briefings in Bioinformatics*, 25(5).
85. Elbashir, M. K., Almotilag, A., Mahmood, M. A., & Mohammed, M. (2024). Enhancing non-small cell lung cancer survival prediction through multiomics integration using graph attention network. *Diagnostics*, 14(19), 2178.
86. Dickinson, Q., Kohler, A., Ott, M., & Meyer, J. G. (2022). Multiomic integration by machine learning (MIMaL). *Bioinformatics*, 38(21), 4908-4918.
87. Li, B., & Nabavi, S. (2024). A multimodal graph neural network framework for cancer molecular subtype classification. *BMC bioinformatics*, 25(1), 27.
88. Liu, Z., & Park, T. (2024). DMOIT: denoised multiomics integration approach based on transformer multihead self-attention mechanism. *Frontiers in Genetics*, 15, 1488683.
89. Waqas, A., Tripathi, A., Ramachandran, R. P., Stewart, P. A., & Rasool, G. (2024). Multimodal data integration for oncology in the era of deep neural networks: a review. *Frontiers in Artificial Intelligence*, 7, 1408843.
90. Kim, M., Park, J., Oh, S., Jeong, B. H., Byun, Y., Shin, S. H., ... & Cho, E. H. (2024). Deep learning model integrating cfDNA methylation and fragment size profiles for lung cancer diagnosis. *Scientific reports*, 14(1), 14797.
91. Newman, A. M., Lovejoy, A. F., Klass, D. M., Kurtz, D. M., Chabon, J. J., Scherer, F., ... & Alizadeh, A. A. (2016). Integrated digital error suppression for improved detection of circulating tumor DNA. *Nature biotechnology*, 34(5), 547-555.

92. Cai, Z., Poulos, R. C., Aref, A., Robinson, P. J., Reddel, R. R., & Zhong, Q. (2024). DeePathNet: A Transformer-Based Deep Learning Model Integrating Multiomic Data with Cancer Pathways. *Cancer Research Communications*, 4(12), 3151-3164.

93. Pati, S., Kumar, S., Varma, A., Edwards, B., Lu, C., Qu, L., ... & Bakas, S. (2024). Privacy preservation for federated learning in health care. *Patterns*, 5(7).

94. Nan, Y., Del Ser, J., Walsh, S., Schönlieb, C., Roberts, M., Selby, I., ... & Yang, G. (2022). Data harmonization for information fusion in digital healthcare: A state-of-the-art systematic review, meta-analysis and future research directions. *Information Fusion*, 82, 99-122.

95. Li, M., Xu, P., Hu, J., Tang, Z., & Yang, G. (2025). From challenges and pitfalls to recommendations and opportunities: Implementing federated learning in healthcare. *Medical Image Analysis*, 103497.

96. Rajendran, S., Pan, W., Sabuncu, M. R., Chen, Y., Zhou, J., & Wang, F. E. I. (2024). Learning across diverse biomedical data modalities and cohorts: Challenges and opportunities for innovation. *Patterns*, 5(2).

97. Al Aziz, M. M., Anjum, M. M., Mohammed, N., & Jiang, X. (2022). Generalized genomic data sharing for differentially private federated learning. *Journal of Biomedical Informatics*, 132, 104113.

98. Froelicher, D., Troncoso-Pastoriza, J. R., Raisaro, J. L., Cuendet, M. A., Sousa, J. S., Cho, H., ... & Hubaux, J. P. (2021). Truly privacy-preserving federated analytics for precision medicine with multiparty homomorphic encryption. *Nature communications*, 12(1), 5910.

99. Kolobkov, D., Mishra Sharma, S., Medvedev, A., Lebedev, M., Kosaretskiy, E., & Vakhitov, R. (2024). Efficacy of federated learning on genomic data: a study on the UK Biobank and the 1000 Genomes Project. *Frontiers in big data*, 7, 1266031.

100. Zhu, Y., Chen, W., Zhu, K., Liu, Y., Huang, S., & Zeng, P. (2024). Polygenic prediction for underrepresented populations through transfer learning by utilizing genetic similarity shared with European populations. *Briefings in Bioinformatics*, 26(1).

101. Lerga-Jaso, J., Terpolovsky, A., Novković, B., Osama, A., Manson, C., Bohn, S., ... & Yazdi, P. G. (2025). Optimization of multiancestry polygenic risk score disease prediction models. *Scientific Reports*, 15(1), 17495.

102. Han, A. L., Sands, C. F., Matelska, D., Butts, J. C., Ravanmehr, V., Hu, F., ... & Dhindsa, R. S. (2025). Diverse ancestral representation improves genetic intolerance metrics. *Nature communications*, 16(1), 2648.

103. Lupi, A. S., Vazquez, A. I., & de Los Campos, G. (2024). Mapping the relative accuracy of cross-ancestry prediction. *Nature Communications*, 15(1), 1-14.

104. Zhu, X., Chu, C. J., Pan, W., Li, Y., Huang, H., & Zhao, L. (2022). The correlation between periodontal parameters and cell-free DNA in the gingival crevicular fluid, saliva, and plasma in Chinese patients: a cross-sectional study. *Journal of Clinical Medicine*, 11(23), 6902.

105. Lachenmeier, D. W., & Monakhova, Y. B. (2011). Short-term salivary acetaldehyde increase due to direct exposure to alcoholic beverages as an additional cancer risk factor beyond ethanol metabolism. *Journal of experimental & clinical cancer research*, 30(1), 3.

106. Zięba, S., Maciejczyk, M., & Zalewska, A. (2022). Ethanol-and cigarette smoke-related alternations in oral redox homeostasis. *Frontiers in Physiology*, 12, 793028.

107. Shreya, S., Annamalai, M., Jirge, V. L., & Sethi, S. (2023). Utility of salivary biomarkers for diagnosis and monitoring the prognosis of nicotine addiction-A systematic review. *Journal of Oral Biology and Craniofacial Research*, 13(6), 740-750.

108. Hu, Y., Xu, M., Liu, M., & Peng, H. (2025). Comparison of saliva and blood derived cell free RNAs for detecting oral squamous cell carcinoma. *Scientific Reports*, 15(1), 4645.

109. Saidak, Z., Milly, M., Louandre, C., Colin, E., Rusu, P. M., Paasche, A., ... & Galmiche, A. (2025). Salivary cell-free DNA methylation analysis for oncological monitoring of surgical resection of oral squamous cell carcinoma. *Frontiers in Oral Health*, 6, 1614371.

110. Elashoff, D., Zhou, H., Reiss, J., Wang, J., Xiao, H., Henson, B., ... & Wong, D. T. (2012). Prevalidation of salivary biomarkers for oral cancer detection. *Cancer Epidemiology, Biomarkers & Prevention*, 21(4), 664-672.

111. Romsos, E. L., Kline, M. C., Duewer, D. L., Toman, B., & Farkas, N. (2018). *Certification of Standard Reference Materials® 2372a human DNA quantitation* (NIST Special Publication 260-189). National Institute of Standards and Technology.
112. Huggett, J. F. (2020). The digital MIQE guidelines update: minimum information for publication of quantitative digital PCR experiments for 2020. *Clinical chemistry*, 66(8), 1012-1029.
113. Sathyanarayana, S. H., Spracklin, S. B., Deharvengt, S. J., Green, D. C., Instasi, M. D., Gallagher, T. L., ... & Tsongalis, G. J. (2025). Standardized Workflow and Analytical Validation of Cell-Free DNA Extraction for Liquid Biopsy Using a Magnetic Bead-Based Cartridge System. *Cells*, 14(14), 1062.
114. Chu, L. J., Chang, Y. T., Chien, C. Y., Chung, H. C., Wu, S. F., Chen, C. J., ... & Yu, J. S. (2024). Clinical validation of a saliva-based matrix metalloproteinase-1 rapid strip test for detection of oral cavity cancer. *Biomedical Journal*, 47(1), 100594.
115. Ghosh, S. K., Man, Y., Fraiwan, A., Waters, C., McKenzie, C., Lu, C., ... & Weinberg, A. (2024). Beta-defensin index: A functional biomarker for oral cancer detection. *Cell Reports Medicine*, 5(3).
116. Nikanjam, M., Kato, S., Allen, T., Sicklick, J. K., & Kurzrock, R. (2025). Novel clinical trial designs emerging from the molecular reclassification of cancer. *CA: A Cancer Journal for Clinicians*, 75(3), 243-267.
117. Vickers, A. J., van Calster, B., & Steyerberg, E. W. (2019). A simple, step-by-step guide to interpreting decision curve analysis. *Diagnostic and prognostic research*, 3(1), 18.
118. Chalkou, K., Vickers, A. J., Pellegrini, F., Manca, A., & Salanti, G. (2023). Decision curve analysis for personalized treatment choice between multiple options. *Medical decision making*, 43(3), 337-349.
119. Raman, S., Shafie, A. A., Tan, B. Y., Abraham, M. T., Chen Kiong, S., & Cheong, S. C. (2023, April). Economic evaluation of oral cancer screening programs: Review of outcomes and study designs. In *Healthcare* (Vol. 11, No. 8, p. 1198). MDPI.
120. Kumdee, C., Kulpeng, W., & Teerawattananon, Y. (2018). Cost-utility analysis of the screening program for early oral cancer detection in Thailand. *PloS one*, 13(11), e0207442.
121. Dwivedi, P., Lohiya, A., Bahuguna, P., Singh, A., Sulaiman, D., Singh, M. K., ... & Rizwan, S. A. (2023). Cost-effectiveness of population-based screening for oral cancer in India: an economic modeling study. *The Lancet Regional Health-Southeast Asia*, 16.
122. AlZoubi, I. A., Alam, M. K., Alkayid, R. B. S., Alruwaili, H. O., Alruwaili, E. O., Alsaleh, T. A., ... & Noor, N. F. B. M. (2025). Evaluation of Salivary Biomarkers in the Early Detection of Oral Squamous Cell Carcinoma. *Journal of Pharmacy and Bioallied Sciences*, 17(Suppl 2), S1276-S1278.
123. Rigotti, P., Polizzi, A., Quinzi, V., Blasi, A., Lombardi, T., Lo Muzio, E., & Isola, G. (2025). Cell-Free DNA as a Prognostic Biomarker in Oral Carcinogenesis and Oral Squamous Cell Carcinoma: A Translational Perspective. *Cancers*, 17(14), 2366.
124. Liu, Z., Zhou, Y., Lu, J., Gong, T., Ibáñez, E., Cifuentes, A., & Lu, W. (2024). Microfluidic biosensors for biomarker detection in body fluids: a key approach for early cancer diagnosis. *Biomarker Research*, 12(1), 153.
125. Salfer, B., Havo, D., Kuppinger, S., Wong, D. T., Li, F., & Zhang, L. (2023). Evaluating preanalytical variables for saliva cell-free DNA liquid biopsy. *Diagnostics*, 13(10), 1665.
126. Hu, C. C., Wang, S. G., Gao, Z., Qing, M. F., Pan, S., Liu, Y. Y., & Li, F. (2025). Emerging salivary biomarkers for early detection of oral squamous cell carcinoma. *World Journal of Clinical Oncology*, 16(4), 103803.
127. Elgendi, M., Lyzwinski, L., Kübler, E., Shokurov, A. V., Howard, N., & Menon, C. (2024). Advancing cancer detection with portable salivary sialic acid testing. *npj Biosensing*, 1(1), 3.
128. Chahartangi, F., Zarifsanaiey, N., Mehrabi, M., & Ghoochani, B. Z. (2025). Integrating augmented reality virtual patients into healthcare training: A scoping review of learning design and technical requirements. *PloS one*, 20(7), e0324740.
129. Kiseleva, O. I., Ikhalaynen, Y. A., Kurbatov, I. Y., Arzumanyan, V. A., Kryukova, P. A., & Poverennaya, E. V. (2025). Dried Spot Paradigm: Problems and Prospects in Proteomics. *International Journal of Molecular Sciences*, 26(8), 3857.
130. Li, Y., Ou, Y., Fan, K., & Liu, G. (2024). Salivary diagnostics: opportunities and challenges. *Theranostics*, 14(18), 6969.

131. Hamidizadeh, M., Martins, R. F., & Bier, F. F. (2025). Point-of-Care Diagnostics Using Self-heating Elements from Smart Food Packaging: Moving Toward Instrument-Free Nucleic Acid-Based Detection. *Molecular Diagnosis & Therapy*, 29(1), 67-80.

132. Rebol, M., Pietroszek, K., Sikka, N., Ranniger, C., Hood, C., Rutenberg, A., ... & Gütl, C. (2023, October). Evaluating Augmented Reality Communication: How Can We Teach Procedural Skill in AR?. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology* (pp. 1-11).

133. Maloney, D., Ong, S. H., Miskovic-Wheatley, J., Dann, K. M., Sidari, M., Hambleton, A., ... & Maguire, S. (2024). The Essentials: Upskilling a National Health Workforce in the Identification and Treatment of Eating Disorders. *International Journal of Eating Disorders*, 57(12), 2427-2437.

134. Okamura, M., Minchin, E., Mazariego, C., Hersch, J., Taylor, N., & Juraskova, I. (2025). Informed Consent for Newborn Genomic Screening: Interest-Holder Perspectives on Dynamic Consent in an Evolving Landscape. *International Journal of Neonatal Screening*, 11(2), 41.

135. Carter-Bawa, L., Valenzona, F., Lucca-Susana, M., Kwok, G., & Shoulders, E. N. (2025). "Paint and Sip for Cancer Prevention": A Novel Arts-Based Community Engagement Strategy to Advance Cancer Education and Screening in Underserved Individuals. *Journal of Cancer Education*, 1-9.

136. Santos Salas, A., Bassah, N., Pujadas Botey, A., Robson, P., Beranek, J., Iyiola, I., & Kennedy, M. (2024). Interventions to improve access to cancer care in underserved populations in high income countries: a systematic review. *Oncology Reviews*, 18, 1427441.

137. Le, L. T. P., Nguyen, A. H. Q., Ngo, H. T. T., Wang, X., Cunningham, B., Valera, E., ... & Do, C. D. (2024). Current smartphone-assisted point-of-care cancer detection: Toward supporting personalized cancer monitoring. *TrAC Trends in Analytical Chemistry*, 174, 117681.

138. Basnayake Ralalage, P., Mitchell, T., Zammit, C., Baynam, G., Kowal, E., Masey, L., ... & Ferdinand, A. (2024). "Equity" in genomic health policies: a review of policies in the international arena. *Frontiers in Public Health*, 12, 1464701.

139. Alsuraifi, A., Mouzan, M. M., Ali, A. a. A., Algzaare, A., Aqeel, Z., Ezzat, D., & Ayad, A. (2025). Revolutionizing Tooth Regeneration: Innovations from Stem Cells to Tissue Engineering. *Regenerative Engineering and Translational Medicine*.

140. Morin, S., Bazarova, N., Jacon, P., & Vella, S. (2018). The manufacturers' perspective on world health organization prequalification of in vitro diagnostics. *Clinical Infectious Diseases*, 66(2), 301-305.

141. Haas, M. A., Madelli, E. O., Brown, R., Prictor, M., & Boughtwood, T. (2024). Evaluation of CTRL: a web application for dynamic consent and engagement with individuals involved in a cardiovascular genetic disorders cohort. *European Journal of Human Genetics*, 32(1), 61-68.

142. Bruns, A., & Winkler, E. C. (2024). Dynamic consent: a royal road to research consent?. *Journal of Medical Ethics*.

143. Miller, D. T., Lee, K., Abul-Husn, N. S., Amendola, L. M., Brothers, K., Chung, W. K., ... & ACMG Secondary Findings Working Group. (2023). ACMG SF v3. 2 list for reporting of secondary findings in clinical exome and genome sequencing: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genetics in Medicine*, 25(8), 100866.

144. Laguna, J. C., Pastor, B., Nalda, I., Hijazo-Pechero, S., Teixido, C., Potrony, M., ... & Mezquita, L. (2024). Incidental pathogenic germline alterations detected through liquid biopsy in patients with solid tumors: prevalence, clinical utility and implications. *British Journal of Cancer*, 130(9), 1420-1431.

145. Batool, S. M., Hsia, T., Beecroft, A., Lewis, B., Ekanayake, E., Rosenfeld, Y., ... & Balaj, L. (2023). Extrinsic and intrinsic preanalytical variables affecting liquid biopsy in cancer. *Cell Reports Medicine*, 4(10).

146. Pantel, K., Alix-Panabières, C., Hofman, P., Stoecklein, N. H., Lu, Y. J., Lianidou, E., ... & Ciliberto, G. (2025). Fostering the implementation of liquid biopsy in clinical practice: meeting report 2024 of the European Liquid Biopsy Society (ELBS). *Journal of Experimental & Clinical Cancer Research*, 44(1), 156.

147. Cui, W., Milner-Watts, C., O'Sullivan, H., Lyons, H., Minchom, A., Bhosle, J., ... & Popat, S. (2022). Upfront cell-free DNA next generation sequencing improves target identification in UK first line advanced non-small cell lung cancer (NSCLC) patients. *European Journal of Cancer*, 171, 44-54.

148. Kim, H., & Park, K. U. (2023). Clinical circulating tumor DNA testing for precision oncology. *Cancer Research and Treatment: Official Journal of Korean Cancer Association*, 55(2), 351-366.

149. Aboy, M., Minssen, T., & Vayena, E. (2024). Navigating the EU AI Act: implications for regulated digital medical products. *NPJ Digital Medicine*, 7(1), 237.

150. Kalodanis, K., Feretzakis, G., Rizomiliotis, P., Verykios, V. S., Papapavlou, C., Skrekas, A., & Anagnostopoulos, D. (2025). Evaluating the Impact of the EU AI Act on Medical Device Regulation. In *Envisioning the Future of Health Informatics and Digital Health* (pp. 40-44). IOS Press.

151. Center for Devices and Radiological Health. (2025, June 26). Cybersecurity in Medical Devices: quality system considerations and content of premarket submissions. U.S. Food And Drug Administration. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/cybersecurity-medical-devices-quality-system-considerations-and-content-premarket-submissions>

152. U.S. Food and Drug Administration. (2025, July 28). Premarket Approval (PMA): Guardant360 CDx (PMA P200010/S008). Retrieved July 30, 2025, from <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P200010S008>

153. Bauml, J. M., Li, B. T., Velcheti, V., Govindan, R., Curioni-Fontecedro, A., Dooms, C., ... & Skoulidis, F. (2022). Clinical validation of Guardant360 CDx as a blood-based companion diagnostic for sotorasib. *Lung cancer*, 166, 270-278.

154. Center for Devices and Radiological Health. (2024, November 7). Breakthrough Devices Program. U.S. Food And Drug Administration. <https://www.fda.gov/medical-devices/how-study-and-market-your-device/breakthrough-devices-program>

155. Cui, W., Milner-Watts, C., McVeigh, T. P., Minchom, A., Bholse, J., Davidson, M., ... & O'Brien, M. (2022). A pilot of Blood-First diagnostic cell free DNA (cfDNA) next-generation sequencing (NGS) in patients with suspected advanced lung cancer. *Lung Cancer*, 165, 34-42.

156. Andersson, D., Kebede, F. T., Escobar, M., Österlund, T., & Ståhlberg, A. (2024). Principles of digital sequencing using unique molecular identifiers. *Molecular Aspects of Medicine*, 96, 101253.

157. Zhang, X., Li, J., Zhuang, Z., Wang, J., Bu, Z., & Lan, X. (2024). Challenges and prospects of cell-free DNA in precision oncology. *Medicine Plus*, 1(4), 100059.

158. Bronkhorst, A. J., & Holdenrieder, S. (2023). Cell-free nucleic acids: physico-chemical properties, analytical considerations, and clinical applications. *Diagnostics*, 13(13), 2312.

159. Hu, Y., Xu, M., Liu, M., & Peng, H. (2025). Comparison of saliva and blood derived cell free RNAs for detecting oral squamous cell carcinoma. *Scientific Reports*, 15(1), 4645.

160. Ahmed, A. A., Sborchia, M., Bye, H., Roman-Escorza, M., Amar, A., Henley-Smith, R., ... & Mathew, C. G. (2024). Mutation detection in saliva from oral cancer patients. *Oral Oncology*, 151, 106717.

161. Saidak, Z., Milly, M., Louandre, C., Colin, E., Rusu, P. M., Paasche, A., ... & Galmiche, A. (2025). Salivary cell-free DNA methylation analysis for oncological monitoring of surgical resection of oral squamous cell carcinoma. *Frontiers in Oral Health*, 6, 1614371.

162. Kumar, P., Gupta, S., & Das, B. C. (2024). Saliva as a potential noninvasive liquid biopsy for early and easy diagnosis/prognosis of head and neck cancer. *Translational oncology*, 40, 101827.

163. Pandey, S., & Yadav, P. (2025). Liquid biopsy in cancer management: Integrating diagnostics and clinical applications. *Practical Laboratory Medicine*, 43, e00446.

164. Rehm, H. L., Page, A. J., Smith, L., Adams, J. B., Alterovitz, G., Babb, L. J., ... & Rodarmer, K. W. (2021). GA4GH: International policies and standards for data sharing across genomic research and healthcare. *Cell genomics*, 1(2).

165. Zhang, J. Y. (2023). Commoning genomic solidarity to improve global health equality. *Cell Genomics*, 3(10).

166. Halmai, N. B., Taitingfong, R., Jennings, L. L., Yracheta, J., Garba, I., Lund, J. R., ... & Carroll, S. R. (2025). Indigenous Data Sovereignty in Genomics and Human Genetics: Genomic Equity and Justice for Indigenous Peoples. *Annual Review of Genomics and Human Genetics*, 26.

167. Roberts, K. J., Ogbagiorgis, W., Sy, A., Williams-Blangero, S., Stewart, L. V., Manusov, E., ... & Madden, E. B. (2025). Increasing the genomic workforce through research capacity building: Designing evaluation plans for maximum impact. *The American Journal of Human Genetics*, 112(5), 967-974.

168. Borea, R., Saldanha, E. F., Maheswaran, S., Nicolo, E., Singhal, S., Pontolillo, L., ... & Reduzzi, C. (2025). Cancer in a drop: Advances in liquid biopsy in 2024. *Critical Reviews in Oncology/Hematology*, 104776.

169. Ma, L., Guo, H., Zhao, Y., Liu, Z., Wang, C., Bu, J., ... & Wei, J. (2024). Liquid biopsy in cancer: current status, challenges and future prospects. *Signal Transduction and Targeted Therapy*, 9(1), 336.
170. Batool, S. M., Yekula, A., Khanna, P., Hsia, T., Gamblin, A. S., Ekanayake, E., ... & Balaj, L. (2023). The Liquid Biopsy Consortium: Challenges and opportunities for early cancer detection and monitoring. *Cell Reports Medicine*, 4(10).
171. Connors, D., Allen, J., Alvarez, J. D., Boyle, J., Cristofanilli, M., Hiller, C., ... & Stewart, M. (2020). International liquid biopsy standardization alliance white paper. *Critical reviews in oncology/hematology*, 156, 103112.
172. McKibbin, K., & Shabani, M. (2023). Genomic data as a national strategic resource: implications for the genomic commons and international data sharing for biomedical research and innovation. *Journal of Law, Medicine & Ethics*, 51(2), 301-313.
173. Van Dyk, L. (2014). A review of telehealth service implementation frameworks. *International journal of environmental research and public health*, 11(2), 1279-1298.
174. Surdu, A., Foia, L. G., Luchian, I., Trifan, D., Tatarciuc, M. S., Scutariu, M. M., ... & Budala, D. G. (2025). Saliva as a Diagnostic Tool for Systemic Diseases—A Narrative Review. *Medicina*, 61(2), 243.
175. Li, Y., Ou, Y., Fan, K., & Liu, G. (2024). Salivary diagnostics: opportunities and challenges. *Theranostics*, 14(18), 6969.
176. Chohan, M. A., Li, T., Ramakrishnan, S., & Sheraz, M. (2025). Artificial Intelligence in Financial Risk Early Warning Systems: A Bibliometric and Thematic Analysis of Emerging Trends and Insights. *International Journal of Advanced Computer Science & Applications*, 16(1).
177. Huang, X., Fayoumi, A., Winter, E., & Najdawi, A. (2025). Design requirements of Breast Cancer Symptom-Management Apps. *Informatics*, 12(3), 72.
178. National Institute of Standards and Technology. (2023, January). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>.
179. Campos, S., Papadatos, H., Roger, F., Touzet, C., Quarks, O., & Murray, M. (2025). A Frontier AI Risk Management Framework: Bridging the Gap Between Current AI Practices and Established Risk Management. *arXiv preprint arXiv:2502.06656*.
180. National Institutes of Health, National Institute of Dental and Craniofacial Research. (2025, July). Salivary diagnostics. <https://www.nidcr.nih.gov/grants-funding/funded-research/research-investments-advances/salivary-diagnostics>
181. Rancea, A., Anghel, I., & Cioara, T. (2024). Edge computing in healthcare: Innovations, opportunities, and challenges. *Future internet*, 16(9), 329.
182. Alshehri, F., & Muhammad, G. (2020). A comprehensive survey of the Internet of Things (IoT) and AI-based smart healthcare. *IEEE access*, 9, 3660-3678.
183. Lotter, W., Hassett, M. J., Schultz, N., Kehl, K. L., Van Allen, E. M., & Cerami, E. (2024). Artificial intelligence in oncology: current landscape, challenges, and future directions. *Cancer discovery*, 14(5), 711-726.
184. Bartolomucci, A., Nobrega, M., Ferrier, T., Dickinson, K., Kaorey, N., Nadeau, A., ... & Burnier, J. V. (2025). Circulating tumor DNA to monitor treatment response in solid tumors and advance precision oncology. *NPJ Precision Oncology*, 9(1), 84.
185. Shree, P., Aggarwal, Y., Kumar, M., Majhee, L., Singh, N. N., Prakash, O., ... & Rai, A. (2024). Saliva based diagnostic prediction of oral squamous cell carcinoma using FTIR spectroscopy. *Indian Journal of Otolaryngology and Head & Neck Surgery*, 76(3), 2282-2289.
186. Hu, Y., Xu, M., Liu, M., & Peng, H. (2025). Comparison of saliva and blood derived cell free RNAs for detecting oral squamous cell carcinoma. *Scientific Reports*, 15(1), 4645.
187. Remori, V., Airoidi, M., Alberio, T., Fasano, M., & Azzì, L. (2024). Prediction of Oral Cancer Biomarkers by Salivary Proteomics Data. *International Journal of Molecular Sciences*, 25(20), 11120.
188. Schneider, G., Kaliappan, A., Joos, N., Dooley, L. M., Shumway, B. S., Chaires, J. B., ... & Garbett, N. C. (2024). Evaluation of Thermal Liquid Biopsy Analysis of Saliva and Blood Plasma Specimens as a Novel Diagnostic Modality in Head and Neck Cancer. *Cancers*, 16(24), 4220.

189. Han, Z., Hu, Y., Lin, X., Cheng, H., Dong, B., Liu, X., ... & Xu, Z. Z. (2025). Systematic analyses uncover robust salivary microbial signatures and host-microbiome perturbations in oral squamous cell carcinoma. *MSystems*, 10(2), e01247-24.
190. Qin, Y., Dong, X., & Li, B. (2025). Salivary miRNAs and cytokines associated with diagnosis and prognosis of oral squamous cell carcinoma. *Frontiers in Cell and Developmental Biology*, 13, 1531016.
191. Hu, C. C., Wang, S. G., Gao, Z., Qing, M. F., Pan, S., Liu, Y. Y., & Li, F. (2025). Emerging salivary biomarkers for early detection of oral squamous cell carcinoma. *World Journal of Clinical Oncology*, 16(4), 103803.
192. Liao, S. C., Peng, J., Mauk, M. G., Awasthi, S., Song, J., Friedman, H., ... & Liu, C. (2016). Smart cup: A minimally instrumented, smartphone-based point-of-care molecular diagnostic device. *Sensors and Actuators B: Chemical*, 229, 232-238.
193. Alsuraifi, A., Sulaiman, Z. M., Mohammed, N. a. R., Mohammed, J., Ali, S. K., Abdualihamaid, Y. H., Husam, F., & Ayad, A. (2024). Explore the most recent developments and upcoming outlooks in the field of dental nanomaterials. *Beni-Suef University Journal of Basic and Applied Sciences*, 13(1).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.