ACCSCIENCE PUBLISHING

Eurasian Journal of Medicine and Oncology

ORIGINAL RESEARCH ARTICLE

Determination of specific salivary biochemical profiles in oral and ovarian cancer using machine learning and SHAP analysis

Elena I. Dyachenko and Lyudmila V. Bel'skaya*

Biochemistry Research Laboratory, Omsk State Pedagogical University, Omsk, Omsk Oblast, Russia

Abstract

Introduction: Saliva contains local and systemic molecular signals that may support non-invasive cancer detection, but the diagnostic value of integrated salivary profiles across anatomically distinct malignancies remains unclear.

Objective: This study evaluates whether multivariable salivary signatures could distinguish controls, ovarian cancer, and oral cavity cancer.

Methods: Saliva samples from 75 controls, 89 patients with ovarian cancer, and 49 patients with oral cavity cancer were analyzed for amino acids and nitrogen metabolites, inorganic ions, cytokines, growth factors, tumor-associated markers, lipid peroxidation products, and Fourier-transform infrared (FTIR) features. Results from regularized multinomial logistic regression, random forest, and XGBoost models were compared using repeated stratified five-fold cross-validation within the training set. Final model performance was then evaluated on a held-out test set using accuracy, Cohen's κ, balanced accuracy, macro-F1, multiclass receiver operating characteristic–area under the curve, and log loss. Permutation testing, bootstrap uncertainty estimation, SHapley Additive exPlanations (SHAP) analysis, and decision curve analysis were additionally performed.

Results: XGBoost showed the strongest performance and correctly classified all ovarian and oral cavity cancer cases in the held-out test set. Permutation testing confirmed that performance exceeded that of randomly assigned class labels. SHAP analysis indicated that mainly cytokine and tumor-associated markers drove ovarian cancer predictions, whereas lipid-related FTIR features dominated oral cavity cancer predictions. Decision curve analysis demonstrated positive net benefit across threshold ranges.

Conclusion: Integrated salivary profiling enabled accurate internal differentiation of local and distant malignancies. The findings support biologically distinct salivary signatures, but external validation is required.

Keywords: Machine learning; Saliva; Cancer; Biochemical profile; Cytokines; Amino acids

***Corresponding author:**
Lyudmila V. Bel'skaya
(belskaya@omgpu.ru)

Citation: Dyachenko EI, Bel'skaya LV. Determination of specific salivary biochemical profiles in oral and ovarian cancer using machine learning and SHAP analysis. *Eurasian J Med Oncol*. doi: 10.36922/EJMO026310388

Received: July 27, 2026

Revised: September 5, 2026

Accepted: September 15, 2026

Published online: October 8, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Malignant transformation is a multilevel and biologically heterogeneous process accompanied by interconnected changes in cellular metabolism, inflammatory and redox regulation, intercellular communication, and extracellular matrix remodeling.[1,2]

Consequently, the molecular profile of biological fluids during tumor progression, as a rule, cannot be adequately characterized by a single biomarker.[3-5] Concentrations of individual analytes often overlap in malignant, benign, and physiological conditions, whereas diagnostically significant information may be contained in a set of interconnected features and the nature of their combined changes.[6-8] Multivariate statistical methods allow simultaneous analysis of correlated parameters, and machine learning algorithms can identify nonlinear dependencies and interfeature interactions that may remain invisible in traditional univariate analysis. In this regard, complex biomarker panels are considered a more promising approach to the differential diagnosis of cancer than assessing individual indicators.[9-12]

The application of artificial intelligence in biomarker research thus offers the opportunity to transform multivariate molecular measurements into clinically interpretable diagnostic models.[13,14] Algorithms such as regularized logistic regression, random forest, and extreme gradient boosting (XGBoost) vary in their ability to model linear effects, complex interactions, and nonlinear decision boundaries.[15] Systematic comparison of these algorithms can help determine whether disease-related information is represented by a relatively simple combination of variables or a more complex biological picture.[16] At the same time, cross-validation, calibration analysis, sensitivity testing, and model interpretation methods are essential, as high apparent accuracy in a limited dataset may reflect overfitting rather than reproducible disease biology.[17,18]

Studies using salivary metatranscriptomics, metabolomics, and autoantibody data have demonstrated that machine learning classifiers can distinguish oral cancer patients from controls and can identify compact combinations of informative salivary features.[19] Proteomic and transcriptomic studies have confirmed that saliva contains stable and measurable molecular signals suitable for biomarker analysis.[20-22] Its composition reflects the secretory activity of the salivary glands and the physiological state of the oral mucosa, and it is also influenced by systemic metabolic, immune, and endocrine signals.[23-25] Thus, saliva may reflect local tissue processes and systemic responses to disease.[26] However, interpreting saliva diagnostically requires a multifactorial approach, since saliva composition is influenced by flow rate, oral inflammation, microbial activity, diet, medications, and preanalytical conditions.[27-29]

Oral cancer is a particularly informative model for developing salivary diagnostic approaches, since the tumor and its microenvironment are in direct contact with oral fluid, which can result in the influx of tumor-associated molecules, products of epithelial damage, and mediators produced by stromal, vascular, and immune cells.[30] Corresponding changes in oral squamous cell carcinoma have been demonstrated at the transcriptomic, proteomic, metabolomic, and spectroscopic levels, and machine learning has confirmed the advantage of multimarker combinations over individual indicators.[20,31] However, the diagnostic potential of saliva is not limited to local tumors: ovarian cancer is a model of a distant malignant process in which the salivary profile can be altered by systemic inflammation, metabolic remodeling, vascular transport, altered salivary gland function, and intercellular communication involving extracellular vesicles.[32,33] The potential for detecting such a systemic signal is supported by studies in which combinations of salivary molecular markers were able to distinguish patients with ovarian cancer from healthy controls.[34-36] However, the limited specificity of individual salivary tumor markers and their overlap with benign conditions highlight the need to integrate disparate biological signatures within multivariate diagnostic models. Against this background, oral cancer and ovarian cancer offer complementary models for exploring the diagnostic information contained in saliva. Oral cancer allows for the assessment of a predominantly local, tumor-associated salivary signature, whereas ovarian cancer offers the opportunity to examine whether a distant malignancy elicits a detectable systemic salivary response. Comparing these diseases within a single multiclass model may also reveal whether tumor types differ in individual biomarkers or in broader combinations of inflammatory, metabolic, mineral, oxidative, and structural alterations.

In this study, we integrated salivary amino acids and nitrogenous metabolites, inorganic ions, cytokines, growth factors, tumor-associated markers, lipid peroxidation products, and Fourier-transform infrared spectral (FTIR) characteristics. Multiclass diagnostic models were developed using logistic regression, random forest, and XGBoost to discriminate among controls, ovarian cancer patients, and oral cancer patients. Model performance and robustness were assessed using resampling, calibration, permutation, and sensitivity analyses, while class-specific SHapley Additive exPlanations (SHAP) interpretation was used to identify variables and directions of change that contributed to each diagnostic category. The aim of the study is to evaluate the diagnostic potential of saliva and identify biologically interpretable combinations of its parameters characteristic of local and distant malignancies.

2. Materials and methods

2.1. Study design

The study included oral cancer patients (n = 49, age 52.3

± 12.9 years), ovarian cancer patients (n = 89, age 55.0 ± 12.6 years), and healthy volunteers (n = 75, age 51.7 ± 12.5 years). Oral and ovarian cancer patients were recruited in the emergency department and healthy volunteers in the blood transfusion department at the Clinical Oncology Dispensary (Omsk, Russia). Inclusion criteria: female gender, age 30–70 years, absence of any previous treatment (surgery, chemotherapy, or radiation therapy) at the time of inclusion in the study, and histological confirmation of the diagnosis. We confirm that the patients studied in this work do not overlap with patients described in our previous publications from the same institution.

The subgroup of oral cancer patients included the following localizations: tongue C01–C02 (n = 7, 14.3%), gingiva C03 (n = 2, 4.1%), floor of the mouth C04 (n = 5, 10.2%), palate C05 (n = 3, 6.1%), salivary glands C07–C08 (n = 4, 8.2%), tonsils C09 (n = 7, 14.3%), oropharynx C10 (n = 10, 20.3%), pyriform sinus C12 (n = 4, 8.2%), and pharynx C13 (n = 7, 14.3%). The distribution of patients by stage was as follows: Stage II (n = 6, 12.2%), III (n = 10, 20.4%), and IV (n = 33, 67.4%).

The study included 89 patients with ovarian cancer. All patients had histologically confirmed ovarian cancer of Stage I (n = 35, 39.3%), II (n = 12, 13.5%), III (n = 34, 38.2%), or IV (n = 8, 9.0%) (low-grade serous carcinoma [n = 11], serous borderline tumor [n = 12], endometrioid carcinoma [n = 13], mucinous carcinoma [n = 7], clear cell carcinoma [n = 5], high-grade serous carcinoma [n = 39], and undifferentiated cancer [n = 2]).

Potential clinical factors that could influence saliva composition were considered during study planning. Diagnostic groups were formed to ensure no statistically significant differences in age (p = 0.4671) and menopausal status (menopausal/non-menopausal = 50%/50%). Patients with poor periodontal and oral health, as well as current smokers, were excluded from the study to minimize the influence of potential confounding factors. No additional statistical procedures were used to account for these factors in group comparisons and/or machine learning models.

Saliva samples were collected once, in which 66 parameters were determined, including the content of 18 amino acids (Arg, Lys, Orn, Tyr, Trp, Phe, His, Leu, Ile, Met, Val, Gln, Pro, Thr, Ser, Ala, Tau, Gly); 11 ions (NH_4^+, K^+, Na^+, Mg^{2+}, Ca^{2+}, Cl^-, NO_2^-, SO_4^{2-}, NO_3^-, F^-, PO_4^{3-}); 14 characteristics of infrared (IR) absorption bands of lipid extracts (1396H, 1396S, 1458H, 1458S, 2853H, 2853S, 2923H, 2923S, 2957H, 2957S, 2923/2957H, 2923/2957S, 1458/1396H, 1458/1396S; H = intensity, S = area); 8 biochemical parameters (protein, urea, uric acid, imidazole compounds, α-amino acids, diene conjugates, triene conjugates, Schiff bases); 12 cytokines (monocyte chemoattractant protein-1 [MCP-1], vascular endothelial growth factor [VEGF], tumor necrosis factor alpha [TNF-α], interleukin-1 beta [IL-1β], IL-2, IL-4, IL-6, IL-8, IL-10, IL-18, interferon alpha [IFN-α], IFN-γ); and 3 tumor markers (cancer antigen 15-3 [CA15-3], CA27.29, mucin-like carcinoma-associated antigen [MCA).

The selection of salivary parameters was based on the study's objective to characterize the integrated molecular profile of saliva rather than focus on a single biomarker or biological pathway. Therefore, the panel included complementary groups of parameters reflecting various aspects of saliva composition and systemic/local biological processes, including inorganic ions, amino acids, nitrogenous metabolites, protein- and metabolite-related indicators, IR spectral characteristics, cytokines, growth factors, and tumor-associated markers. These parameters were selected based on their presence in the collected saliva samples and their potential biological relevance for metabolic, inflammatory, immune, spectral, and tumor-associated changes described in the literature.

2.2. Collection, storage, and pre-processing of saliva samples

Unstimulated saliva samples were collected in the morning, between 8:00 a.m. and 10:00 a.m., through passive spitting into sterile polypropylene screw-cap centrifuge tubes to a volume of 2 mL. The saliva sample collection procedure included an empty stomach (after an overnight fast), preliminary rinsing of the mouth with distilled water, cessation of smoking for at least two hours, and abstinence from alcohol for at least two days. All volunteers were examined by a dentist for active caries, acute periodontitis, and inflammatory or purulent processes in the oral cavity. If any of these factors were present, volunteers were excluded before saliva sample collection. Saliva samples of insufficient quality or volume, or containing blood, were also not used.

Immediately after collection, the samples were centrifuged at 10,000 × g for 10 min using a CLb-16 centrifuge (Changsha Xiangzhi Centrifuge Instrument Co., Ltd., China) to remove cellular debris and reduce turbidity. The resulting supernatant was transferred to Eppendorf tubes and stored at −80 °C until analysis.

2.3. Determination of salivary composition

Capillary electrophoresis (KAPEL-105M, Lumex, Russia) was used to determine salivary cations (NH_4^+, K^+, Na^+, Mg^{2+}, Ca^{2+}), anions (Cl^-, NO_2^-, SO_4^{2-}, NO_3^-, F^-, PO_4^{3-}), and amino acids.[37] After preliminary extraction with Folch solution (chloroform:ethanol = 2:1, vol.), the intensities (H) and areas (S) for the absorption bands of salivary

lipids were determined using IR spectroscopy (FT-801, Simex, Russia).[38] The absorption bands at 1,396 (δCH_3), 1,458 (δCH_2), 2,853 ($\nu_s CH_3$), 2,923 ($\nu_{as} CH_2$) and 2,957 cm^{-1} ($\nu_{as} CH_3$) were analyzed, as well as the ratios 2,923/2,957 and 1,458/1,396 (ν_s: symmetric stretching; ν_{as}: asymmetric stretching; δ: deformation vibrations).[38]

Salivary concentrations of IL-1β (A-8766), IL-2 (A-8772), IL-4 (A-8754), IL-6 (A-8768), IL-8 (A-8762), IL-10 (A-8774), IL-18 (A-8770), IFN-α (A-8758), IFN-γ (A-8752), MCP-1 (A-8782), TNF-α (A-8756), and VEGF (A-8784) were measured using Vector-Best kits (Russia) with Multiskan FC microplate reader (Thermo Fisher Scientific, United States of America [USA]). A 100 μL aliquot of each saliva sample was used for all assays. Cytokine concentrations were expressed in pg/mL, whereas VEGF concentrations were reported in mU/mL.

In each saliva sample, total protein was quantified using the pyrogallol red method (mg/L; B-8084, Vector-Best), urea was measured using the urease–salicylate Berthelot method (mmol/L; B-8074, Vector-Best), and uric acid was determined using the uricase method (UA, μmol/l; B-8098, Vector-Best). The total concentration of α-amino acids (α-AAs, mmol/L) was determined using the ninhydrin reaction, and imidazole compounds (ICs, mmol/L) were assessed with a diazotization reaction in the presence of sulfanilic acid. All measurements were performed using a StatFax 3300 semi-automated biochemical analyzer (Awareness Technology, USA).[39] The content of lipid peroxidation products (diene conjugates [DC], triene conjugates [TC], Schiff bases [SB]; conventional unit) was determined using the Volchegorsky method.[40]

Technical errors were eliminated by running multiple parallel assays in each case (two parallel plates for enzyme-linked immunosorbent assay [ELISA]; three aliquots of each sample for capillary electrophoresis, biochemical analysis, and FTIR spectroscopy; all sample preparation steps performed separately for each sample). Samples were randomly assigned to batches for ELISA; the remaining assays were performed without division into batches.

The separate assessment of 2957H was motivated by its disproportionately large contribution to model predictions and substantial overlap with information represented by other spectral features. Therefore, its exclusion was considered a sensitivity analysis to assess whether model performance and predictor patterns were robust to a potentially dominant and partially redundant feature, rather than because the feature was considered biologically invalid. This analysis was performed in addition to assessing extreme 2957H values within the oral cancer group.

2.4. Statistical data processing and machine learning model building

Statistical processing and machine learning model building were performed in the R software environment (version 4.4.2; R Foundation for Statistical Computing, Austria) using the tidymodels, glmnet, ranger, xgboost, yardstick, shapviz, dplyr, and ggplot2 packages. To ensure reproducibility, the initial value of the random number generator was fixed (seed = 123). Quantitative indicators were presented as the median and interquartile range (IQR) (Me [Q1; Q3]). Between-group differences among three independent groups were assessed using the nonparametric Kruskal–Wallis test. For pairwise comparisons of groups, Dunn's post hoc test with Bonferroni correction for multiple comparisons was used. The magnitude of the between-group effect was estimated using the ε^2 (epsilon squared) coefficient. All statistical tests were two-sided; differences were considered statistically significant at adjusted $p < 0.05$. The analysis of each indicator was performed based on available observations without imputation of missing values; if the actual sample size differed from the total number of participants, the corresponding *n* value was indicated in the table.

A multiclass classification problem with three outcomes was solved: control, ovarian cancer, and oral cancer. Patient identifiers, age, and service variables directly encoding the diagnosis were excluded from the set of predictors. In a separate scenario, 2957H values below Q1−1.5 × IQR or above Q3+1.5 × IQR in the oral cancer group were considered outliers and excluded; a modified *z*-test with a threshold of $|z| > 3.5$ was used as an additional diagnostic check, while the final exclusion decision was based on the IQR criterion. The data were stratified into training and test sets at a ratio of 75:25, maintaining the proportions of diagnostic classes. Missing values for numeric predictors were replaced with the median, and for categorical predictors, with the mode. Features with zero variance were excluded, and categorical variables, if present, were transformed into indicator variables. For regularized logistic regression, numerical predictors were additionally standardized. All preprocessing parameters were determined only on the training data and within the resampling procedure. Results from regularized multinomial logistic regression, random forest, and XGBoost were compared. Hyperparameter selection was performed using five-fold stratified cross-validation with five iterations, i.e., 25 splits of the training set. For each candidate hyperparameter combination, model performance was evaluated across all resamples of repeated stratified five-fold cross-validation performed

exclusively within the training set (five folds repeated five times). The final hyperparameter combination was selected according to the lowest mean multinomial log loss across all cross-validation resamples, rather than accuracy or F1 score. The selected hyperparameters were then used to finalize the model, which was subsequently evaluated once on the held-out test set that was not used during hyperparameter tuning. Optimal parameters were selected based on the minimum multinomial log loss. Model performance on the held-out test set was assessed using accuracy, Cohen's κ coefficient, and multinomial log loss, and confusion matrices were analyzed. For the extended XGBoost evaluation, balanced accuracy, macro-F1, macro-sensitivity, macro-specificity, and multi-class receiver operating characteristic–area under the curve (ROC–AUC) were calculated using the Hand–Till method.

The robustness of the XGBoost model and the dependence of its results on the 2957H feature were assessed in a pairwise sensitivity analysis. Three scenarios were compared using the same patients and identical splits of repeated cross-validation: a model with all predictors, a model without 2957H, and a single-factor model including only 2957H. For each scenario, mean metric values and empirical intervals corresponding to the 2.5th and 97.5th percentiles of estimates on individual splits were calculated. Predictions obtained for each patient outside the training subset were averaged across cross-validation repetitions. Uncertainty in the 2957H-less model was estimated using a stratified bootstrap at the patient level with 2,000 replicates and the calculation of 95% percentile intervals; a multiclass Brier score was additionally determined. The probability of obtaining high performance due to random data structure was tested using a permutation test: diagnostic labels were shuffled across training subsets, the model was retrained, and the null distributions of the metrics were compared with the results of the original model. Predictions were interpreted using class-specific SHAP values; the global significance of features was determined with the mean absolute SHAP value, and the direction of the relationship between the initial value of a predictor and its SHAP contribution was determined using the Spearman correlation coefficient. In this study, SHAP reflected statistical relationships within the model rather than causal biology. The stability of the predictor ranking was assessed in all 25 splits of the repeated cross-validation using the frequency of the feature in the top 10, median rank, interquartile range of the rank, and reproducibility of the direction of the SHAP effect. Correlations between spectral features were calculated using Spearman's method for the entire sample and separately for each diagnostic group; *p*-values were adjusted using the Benjamini–Hochberg method, and associations with $|\rho| \geq 0.70$ were considered strong. The potential clinical utility of the model was assessed using decision curve analysis (DCA) with a "one-class versus the rest" design by comparing the net utility of the model with the treat-all and treat-none strategies across a threshold probability range of 0.01–0.80. All of these procedures were considered internal validation; external validation on an independent cohort was not performed.

3. Results

3.1. Performance of classification models before and after excluding predictor 2957H

In the baseline scenario, which included all predictors, including 2957H, all three models demonstrated high classification accuracy on the test set (Figure 1A). Accuracy values were 0.982 for random forest, 0.964 for XGBoost, and 0.927 for regularized logistic regression. A similar pattern was observed for Cohen's κ: the highest value was obtained for random forest (0.972), followed by XGBoost (0.944) and regularized logistic regression (0.888). The minimum multinomial log loss was also observed for random forest (0.107), while XGBoost and regularized logistic regression had 0.113 and 0.170, respectively. Thus, in the initial scenario, the most favorable combination of classification accuracy and quality of probabilistic predictions was obtained for random forest, while the results of XGBoost were similar (Figure 1A). After removing the 2957H outliers in the oral cancer group and excluding the 2957H predictor, the models retained high classification ability (Figure 1C). Accuracy was 0.945 for random forest and XGBoost and 0.927 for regularized logistic regression. Cohen's κ values were the same for random forest and XGBoost (0.917), while for regularized logistic regression, it was 0.888. Moreover, XGBoost was characterized by the lowest multinomial log loss (0.125), compared with 0.179 for regularized logistic regression and 0.185 for random forest. Therefore, after excluding 2957H, random forest and XGBoost showed comparable classification accuracy, but XGBoost provided better probabilistic predictions (Figure 1C).

The confusion matrix of the XGBoost model in the baseline scenario confirmed high classification performance (Figure 1B). All oral cancer samples (13 of 13 [100%]) and ovarian cancer samples (23 of 23 [100%]) were correctly classified into their respective classes. In the control group, 17 of 19 samples (89.5%) were correctly classified, while two samples (10.5%) were incorrectly classified as oral cancer (Figure 1B).

After removing outliers and excluding 2957H, the model error structure remained virtually unchanged (Figure 1D). All 13 oral cancer and all 23 ovarian cancer

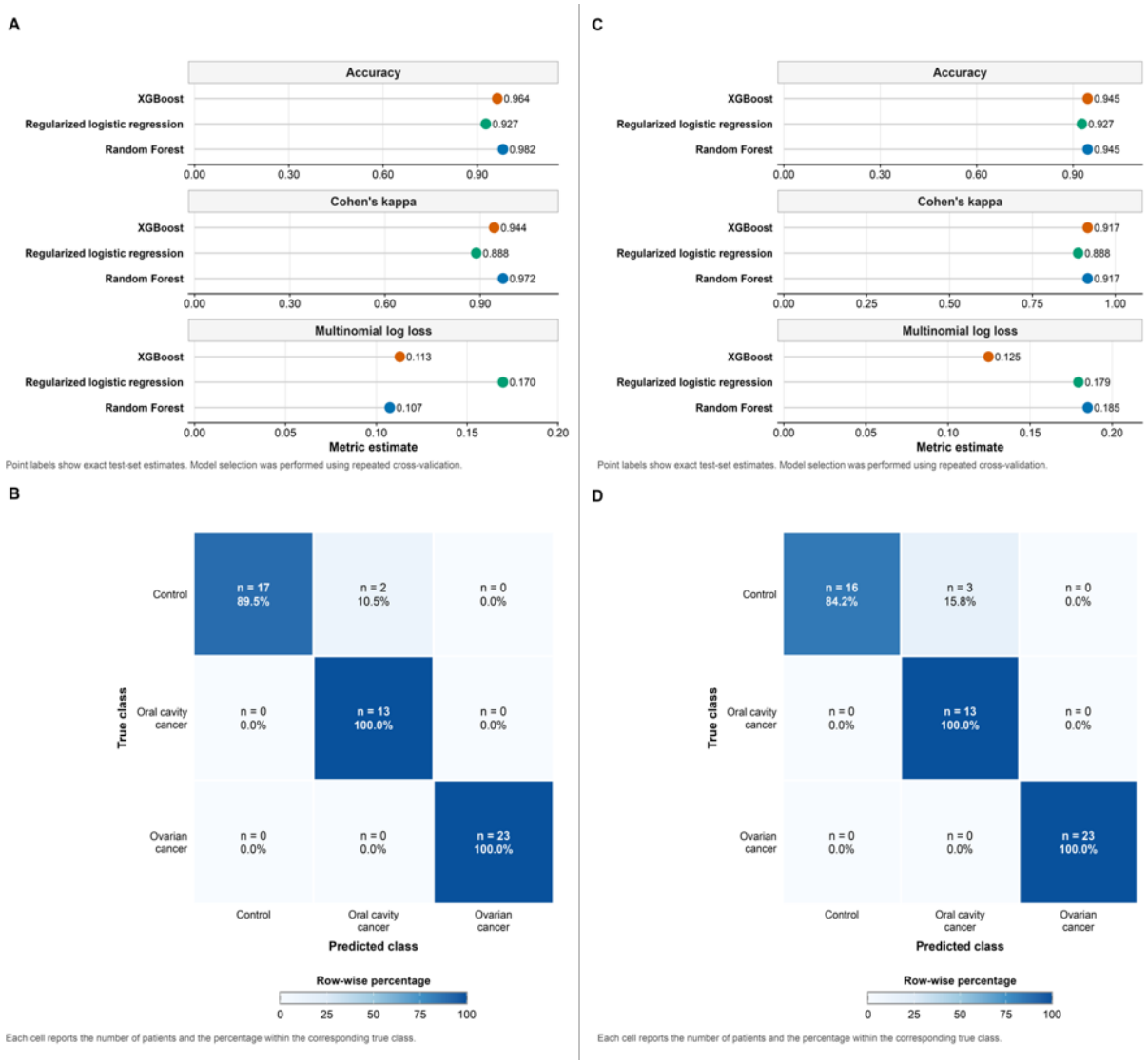


Figure 1. Comparative performance of machine-learning models and extreme gradient boosting (XGBoost) classification results with and without the 2957H predictor. (A & C) Held-out test-set estimates of accuracy, Cohen's κ, and multinomial log loss for XGBoost, regularized multinomial logistic regression, and random forest in models (A) including and (C) excluding 2957H. (B & D) The corresponding row-normalized confusion matrices for XGBoost; excluding 2957H resulted in only a modest decrease in accuracy and Cohen's κ, while both cancer classes remained correctly classified and additional errors were confined to control participants classified as having oral cavity cancer.

samples were correctly classified. In the control group, the number of correctly classified samples decreased to 16 of 19 (84.2%), and three samples (15.8%) were misclassified as oral cancer (Figure 1D). Thus, excluding 2957H did not affect the ability of the XGBoost model to accurately recognize both cancer groups in the test set (Figure 1B and 1D), but it was accompanied by a slight decrease in the proportion of correctly classified control samples. The direction of the errors remained unchanged: misclassified control samples belonged to the oral cancer class, which may reflect partial similarity in their model spectral–biochemical profiles. Given the small size of the test sample, the difference between 89.5% and 84.2% should be interpreted descriptively.

To evaluate the dependence of XGBoost classification performance on the spectral metric 2957H, a sensitivity analysis was conducted using identical repeated stratified cross-validation splits (Figure 2). A single-factor model including only 2957H, a multivariable model excluding 2957H, and a full model with all metrics were compared.

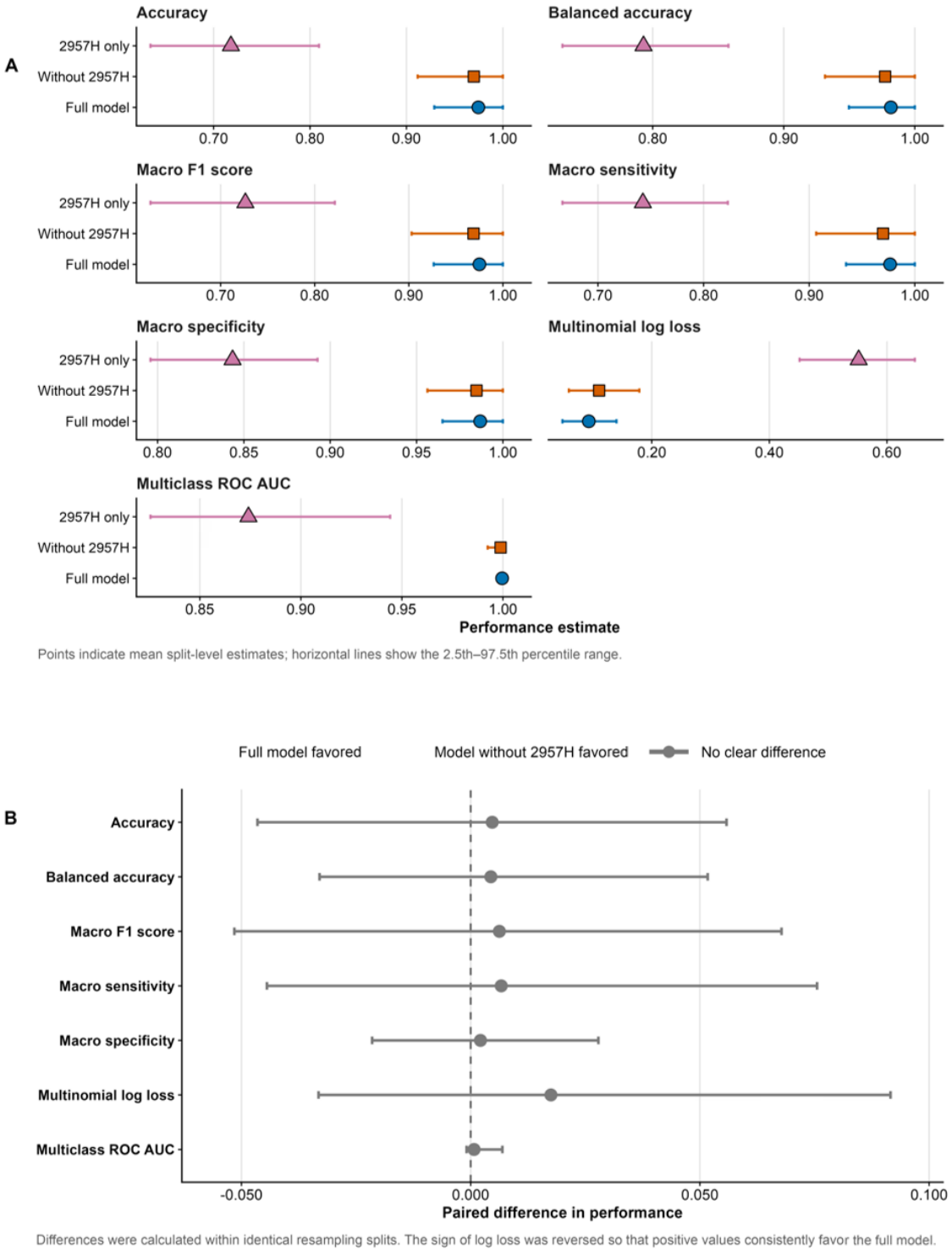


Figure 2. Sensitivity of extreme gradient boosting (XGBoost) classification performance to the inclusion of the 2957H predictor. (A) Comparison between repeated cross-validation estimates and empirical 95% intervals for the full model, the model excluding 2957H, and the single-predictor model based on 2957H alone. Identical patients and resampling splits were used in all scenarios. (B) Paired changes were calculated within identical resampling splits; the small differences between the full model and the model without 2957H, together with intervals crossing zero, indicate no clear deterioration in performance after exclusion of this predictor. Positive values favor the full model; negative values favor the model without 2957H.

The model based only on 2957H demonstrated moderate ability to separate the three diagnostic groups: balanced accuracy was 0.793, macro-F1 was 0.726, multiclass ROC–AUC was 0.874, and multinomial log loss was 0.552. Thus, 2957H had independent classification informativeness, but did not provide performance comparable to multivariable models. Excluding 2957H had virtually no effect on the classification performance of XGBoost. The model without this indicator was characterized by a balanced accuracy of 0.977, macro-F1 of 0.969, multiclass ROC–AUC of 0.9989, and a log loss of 0.110. For the full model, the corresponding indicators were 0.982, 0.975, 0.9996, and 0.093. Thus, the inclusion of 2957H was accompanied by an increase in balanced accuracy by 0.0044, macro-F1 by 0.0063, and multiclass ROC–AUC by 0.0007, as well as a decrease in log loss by 0.0175 (Table 1).

In pairwise comparisons on identical cross-validation splits, mean differences for most metrics favored the full model; however, the magnitudes were small, and percentile intervals included zero. Therefore, excluding 2957H did not systematically reduce classification performance. These results indicate that 2957H provides additional diagnostic information but is not a critical predictor, as its contribution is largely offset by other metrics. To identify potentially interchangeable features, group-specific correlation analysis of spectral metrics was additionally conducted, and the redistribution of SHAP importance after excluding 2957H was assessed.

A patient-level stratified bootstrap analysis confirmed the robustness of the XGBoost model results constructed without the 2957H metric. For the averaged out-of-fold predictions, the balanced accuracy was 0.977, the macro-F1 was 0.969, and the multiclass ROC–AUC was 0.9989; the multinomial log loss was 0.110, and the multiclass Brier score remained low. Narrow 95% bootstrap intervals for the main classification metrics indicated low uncertainty in the estimates and preserved high model performance with repeated resampling of patients (Figure 3).

Thus, excluding 2957H did not prevent the robust separation of the control group, oral cancer patients, and ovarian cancer patients in the internal validation. Permutation validation, which included 1,000 random permutations of diagnostic classes, demonstrated a significant separation between the performance of the real XGBoost model and the null metric distributions. The actual accuracy, balanced accuracy, macro-F1, macro-sensitivity, macro-specificity, and multiclass ROCAUC values significantly exceeded those obtained after class permutation, while the multinomial log loss and multiclass Brier score were significantly lower than the corresponding permutation distributions. In none of the 1,000 permutations did the model performance reach the observed level, corresponding to an empirical $p < 0.001$ and indicating that the high classification performance of the model without 2957H was not due to random assignment of predictors to diagnostic classes.

3.2. Class-specific predictive patterns identified with SHAP analysis

Class-specific SHAP analysis of the XGBoost model built without the 2957H feature revealed different patterns of predictive power for the control group, oral cancer, and ovarian cancer (Figure 4). For the control group, the greatest contribution to the model prediction was IL-6 (mean absolute SHAP value = 0.613), followed by IL-10 (0.379), α-amino acids (0.344), and VEGF (0.342). Higher IL-6 values were predominantly associated with an increase in model output for the control class ($\rho = 0.727$), while increases in IL-10, α-amino acids, and VEGF were accompanied by

Table 1. Repeated cross-validation performance of XGBoost under alternative 2957H inclusion scenarios

Model	Balanced accuracy	Macro-F1	Multiclass ROC–AUC	Log loss
Single-predictor model: 2957H only	0.793 (0.731–0.858)	0.726 (0.626–0.821)	0.8740 (0.8260–0.9440)	0.5520 (0.4510–0.6470)
Multivariable model excluding 2957H	0.977 (0.931–1.000)	0.969 (0.903–1.000)	0.9989 (0.9920–1.0000)	0.1100 (0.0590–0.1789)
Full multivariable model including 2957H	0.982 (0.949–1.000)	0.975 (0.926–1.000)	0.9996 (0.9970–1.0000)	0.0930 (0.0480–0.1399)

Notes: Values are presented as mean estimates with empirical 95% intervals from repeated stratified cross-validation splits. Higher values indicate better performance for balanced accuracy, macro-F1, and multiclass receiver operating characteristic–area under the curve (ROC–AUC), whereas lower values indicate better performance for multinomial log loss. Removal of 2957H produced only minor changes in balanced accuracy, macro-F1, multiclass ROC-AUC, and multinomial log loss, whereas the model based exclusively on 2957H demonstrated substantially poorer performance.
Abbreviation: XGBoost: Extreme gradient boosting.

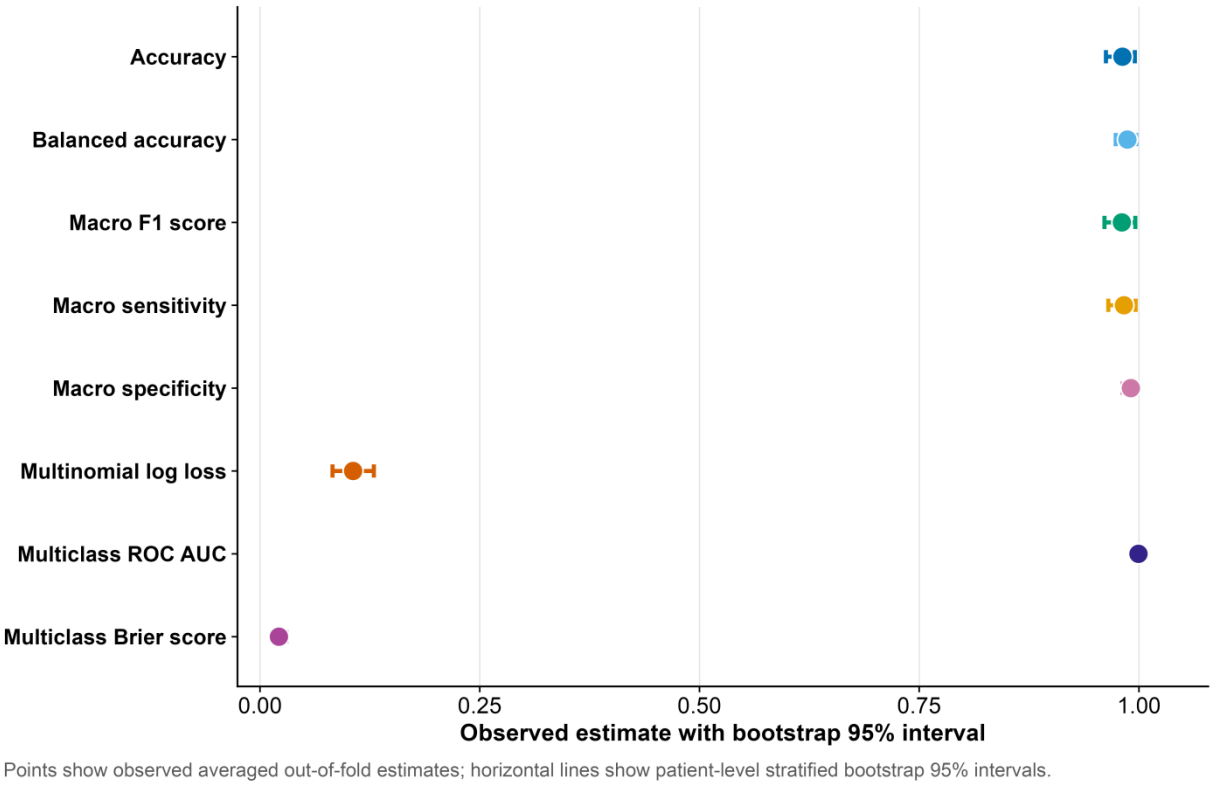


Figure 3. Bootstrap uncertainty of averaged out-of-fold extreme gradient boosting (XGBoost) performance after excluding 2957H. The figure presents patient-level averaged out-of-fold performance estimates for the XGBoost model excluding 2957H, together with stratified bootstrap 95% confidence intervals. The consistently high classification metrics and low multinomial log loss and Brier score indicate stable internal performance after excluding 2957H, although these estimates do not replace validation in an independent cohort.

a decrease (ρ = −0.676, −0.953, and −0.850, respectively). For ovarian cancer, the leading predictors were IL-6 (mean absolute SHAP value = 1.764), TNF-α (0.573), and MCP-1 (0.238). An increase in TNF-α and MCP-1 concentrations was associated with an increase in the model yield for this class (ρ = 0.717 and 0.957, respectively), while an increase in IL-6 was mainly accompanied by a decrease (ρ = −0.880). In the classification of oral cancer, the spectral indices 2923H, 2853S, 1458H, and 2957S had the greatest significance, with mean absolute SHAP values of 1.547, 0.239, 0.139, and 0.136, respectively. For the specified features, a negative correlation was found between their initial values and SHAP contribution (ρ = −0.571, −0.443, −0.918, and −0.720), indicating that a predominant increase in the contribution to the prognosis of oral cancer is associated with a decrease in the corresponding spectral characteristics.

After excluding 2957H, the most pronounced redistribution of model significance was observed in predicting oral cancer. 2923H moved from 21st to first position, and its average absolute SHAP value increased from virtually 0 to 1.547. 2853S moved from 20th to second position, reaching an average absolute SHAP value of 0.239. Increases in relative significance were also noted for 2957S (moved from 23rd to fourth), 2923S (22nd to sixth), and 2853H (19th to seventh). The stability of this redistribution was confirmed with the results of repeated cross-validation: 2923H was among the 10 most significant predictors in all 25 partitions, characterized by a median rank of 1 and the absence of interquartile variability in rank; the direction of its association with the SHAP contribution was preserved in all partitions. The indicators 2853H, 1458H, 2853S, and 2957S were among the 10 most significant predictors in 96%, 88%, 84%, and 84% of partitions, respectively. In the ovarian cancer model, the composition of the leading predictors did not change significantly, whereas the increase in the position of 2923H in the control class (from 33rd to 8th) was less reproducible, since this indicator was among the 10 most significant features in only 28% of partitions. The results indicate that preserving high classification efficiency after excluding 2957H was associated primarily with a stable

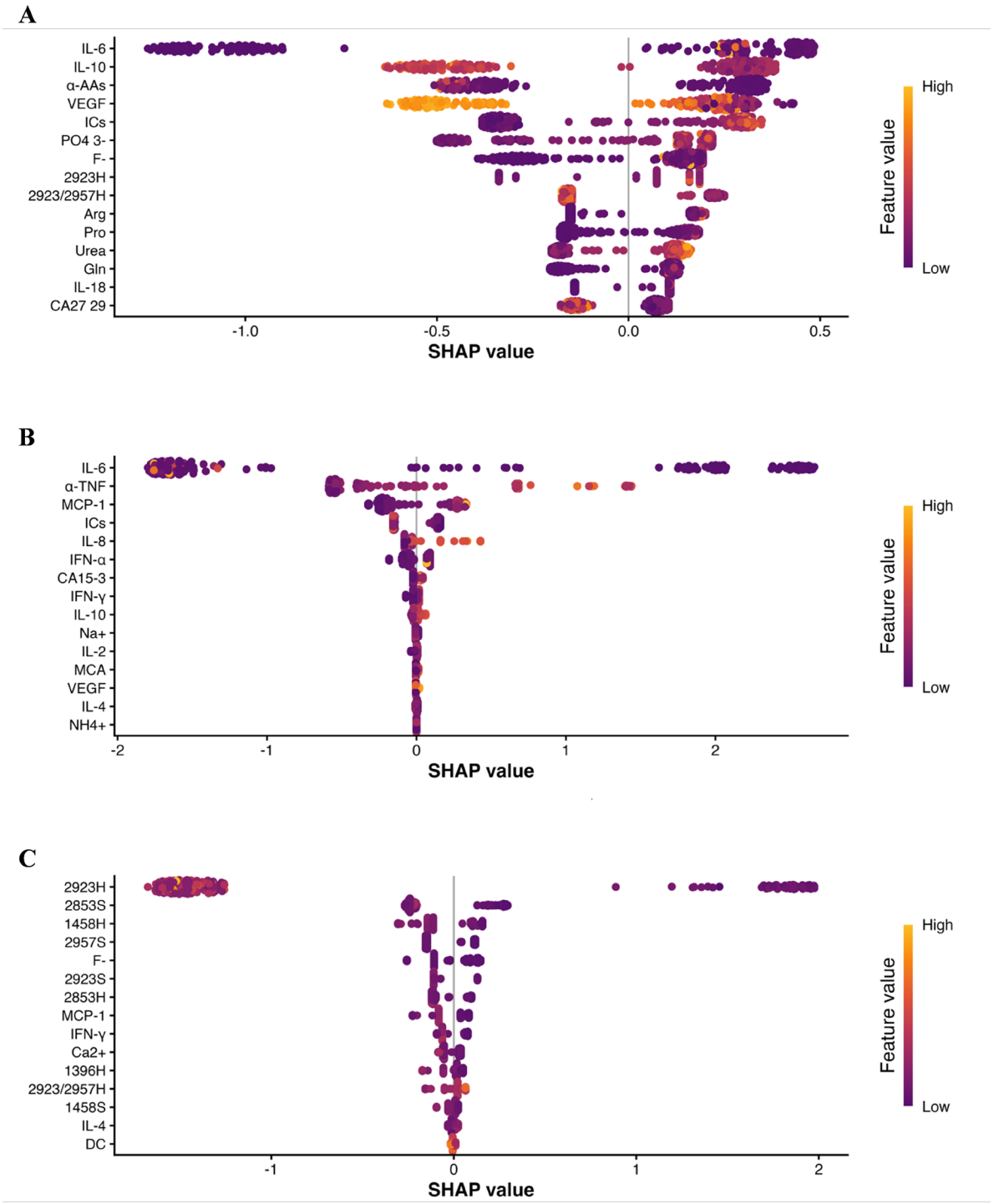


Figure 4. Class-specific SHapley Additive exPlanations (SHAP) profiles of the extreme gradient boosting (XGBoost) model excluding 2957H. Predictors are ordered by class-specific SHAP importance, with positive SHAP values increasing and negative values decreasing the model output for the corresponding diagnostic class. (A) Control predictions were primarily associated with interleukin-6 (IL-6), IL-10, α-amino acids, vascular endothelial growth factor (VEGF), and imidazole compounds (ICs); (B) ovarian cancer predictions were associated with IL-6, tumor necrosis factor (TNF-α), monocyte chemoattractant protein-1 (MCP-1), ICs, and IL-8; and (C) oral cavity cancer predictions were predominantly associated with Fourier transform infrared spectroscopy (FTIR)-derived features, including 2923H, 2853S, 1458H, and 2957S.

redistribution of model significance toward other spectral characteristics in predicting oral cancer.

Table 2 demonstrates that the three diagnostic groups differed not by a single salivary analyte but by coordinated alterations across immunological, metabolic, ionic, oxidative, and FTIR-derived domains. Ovarian cancer was characterized by increased concentrations of IL-10, MCP-1, IFN-α, IFN-γ, TNF-α, VEGF, CA15-3, and MCA, together with lower IL-6 and imidazole compound levels compared with controls, indicating a distinct systemic inflammatory and tumor-associated salivary profile. In contrast, oral cavity cancer was associated with pronounced reductions in proline, glutamine, urea, Ca^{2+}, F^{-}, PO_4^{3-}, diene conjugates, and most FTIR-derived lipid-band intensities, whereas IL-10, VEGF, CA27.29, and the 2923/2957H ratio were increased relative to controls. The largest between-group effects were observed for IL-6 ($\varepsilon^2 = 0.723$), the 2853H, 2853S, 2923H, and 2923S spectral features ($\varepsilon^2 = 0.613–0.633$), IFN-γ ($\varepsilon^2 = 0.507$), and TNF-α ($\varepsilon^2 = 0.447$), suggesting that cytokine alterations and FTIR-derived lipid characteristics contributed most strongly to group differentiation. Here, ε^2 (epsilon squared) represents the effect size associated with the Kruskal–Wallis test and quantifies the magnitude of between-group differences, that is, the proportion of variability in the ranked outcome attributable to group membership. Unlike partial η^2, which is primarily used in parametric ANOVA-type models, ε^2 is appropriate for rank-based non-parametric comparisons such as the Kruskal–Wallis test. Values of approximately 0.01, 0.06, and 0.14 are conventionally interpreted as small, medium, and large effects, respectively. Thus, the observed ε^2 values indicate predominantly large between-group effects. At the same time, the overlapping distributions of several individual markers support using integrated multivariable models rather than isolated biomarkers for diagnostic classification. The biomarkers listed in Table 2 were included in the final classifier.

3.3. Clinical relevance of the XGBoost classification model based on decision curve analysis

Decision curve analysis demonstrated consistently positive clinical benefit of the XGBoost model across the estimated range of threshold probabilities for all three one-versus-all classification tasks (Figure 5).

In DCA, the "treat-all" and "treat-none" strategies are reference decision strategies and should not be interpreted as literal patient treatment. In a real diagnostic setting, "treat all" means classifying every individual as belonging to the diagnostic class in question, while "treat none" means classifying no individuals as belonging to that class. For example, in an analysis of one versus none ovarian cancer, the "treat-all" strategy assumes that all subjects are classified as ovarian cancer cases, while the "treat-none" strategy assumes that none are classified as ovarian cancer cases. These two extreme strategies provide reference levels against which the net benefit of the XGBoost model is compared. A model is considered clinically useful within a given threshold range when its net benefit exceeds the benefit of both reference strategies, indicating a more favorable balance between correctly identified ovarian cancer cases and false-positive classifications. Thus, DCA does not assess whether all patients should actually receive treatment. Rather, it evaluates whether using a model to guide classification leads to better decision-making than two hypothetical extremes: assuming everyone has the disease or assuming some do not.

For the control group, the model maintained a positive and relatively stable net benefit, while the "treat-all" strategy gradually declined, approaching zero at moderate threshold probabilities and becoming negative at higher thresholds. A similar, but more pronounced, pattern was observed for oral cancer. The net benefit of the "treat-all" strategy rapidly declined as the threshold probability increased, while the XGBoost model remained superior to both benchmark strategies throughout almost the entire evaluated range. For ovarian cancer, the model also demonstrated a stable positive net benefit, while the "treat-all" strategy crossed the zero net benefit line approximately midway through the evaluated threshold range and subsequently became increasingly unfavorable. Thus, in all three diagnostic classes, model-guided classification provided a more favorable balance between correctly identified cases and false positive decisions than either classifying all individuals as positive or classifying no individuals as positive. The difference between the model and the all-case treatment strategy became particularly evident at moderate and high threshold probabilities, where false positive classifications receive a gradually increasing penalty. These results demonstrate that XGBoost's discriminatory power translates into potential clinical benefit, rather than merely representing statistical discrimination. The apparent increase or fluctuation in net benefit at the highest threshold probabilities, particularly for the oral cancer class, should be interpreted with caution, as estimates at extreme thresholds may be less stable due to fewer observations influencing clinical decisions.

In all three panels, the solid line corresponding to the model is located above the two baseline strategies: "treat all" and "treat none." This means the model yields a greater net benefit than classifying all patients as positive or, conversely, classifying no patients as positive. For the oral cancer class, the model also demonstrates a positive

Table 2. Distribution and non-parametric comparison of salivary biomarkers among controls and patients with ovarian or oral cancer

Predictor	Control (*n* = 75)	Ovarian cancer (*n* = 89)	Oral cavity cancer (*n* = 49)	Global Kruskal–Wallis test		*p*-value (Dunn post hoc comparisons; Bonferroni-adjusted)		
				p-value (Bonferroni-adjusted)	ε^2	Control vs. ovarian cancer (OVC)	Control vs. oral cavity cancer (OCC)	OVC vs. OCC
Cytokines and growth factors								
IL-2 (pg/mL)	12.17 (9.54–14.11)	10.75 (9.00–13.13)	8.03 (6.58–10.53)	0.0004	0.117	1.000 (ns)	<0.0001 (OCC < control)	0.0001 (OCC < OVC)
IL-4 (pg/mL)	3.62 (2.77–5.83)	3.16 (2.55–4.15)	4.59 (3.06–5.82)	0.0069	0.086	0.0768 (ns)	0.3788 (ns)	0.0002 (OCC > OVC)
IL-6 (pg/mL)	13.48 (12.22–16.94)	3.92 (2.99–5.44)	19.78 (14.27–77.98)	<0.0001	0.723	<0.0001 (OVC < Control)	0.1992 (ns)	<0.0001 (OCC > OVC)
IL-8 (pg/mL)	66.26 (11.56–108.82)	95.00 (42.95–147.41)	61.89 (10.28–130.90)	0.0831	0.058	0.0075 (OVC > Control)	1.0000 (ns)	0.0260 (OCC < OVC)
IL-10 (pg/mL)	8.96 (6.40–14.90)	16.97 (13.42–20.26)	16.81 (12.75–18.99)	<0.0001	0.155	<0.0001 (OVC > Control)	0.0002 (OCC > Control)	1.0000 (ns)
IL-18 (pg/mL)	61.27 (35.49–108.77)	53.25 (25.33–119.71)	33.25 (18.04–81.79)	0.6041	0.035	1.0000 (ns)	0.0247 (OCC < Control)	0.0508 (ns)
MCP-1 (pg/mL)	78.46 (31.54–191.35); *n* = 44	274.62 (165.39–432.31); *n* = 89	63.13 (45.00–83.75); *n* = 45	<0.0001	0.356	<0.0001 (OVC > Control)	0.1232 (ns)	<0.0001 (OCC < OVC)
IFN-α (pg/mL)	9.46 (6.21–12.81)	13.39 (9.29–15.54)	6.61 (4.11–8.93)	<0.0001	0.262	0.0003 (OVC > Control)	0.0585 (ns)	<0.0001 (OCC < OVC)
IFN-γ (pg/mL)	36.52 (23.93–55.32)	90.46 (59.63–136.71)	19.56 (15.78–25.33)	<0.0001	0.507	<0.0001 (OVC > Control)	0.0008 (OCC < Control)	<0.0001 (OCC < OVC)
TNF-α (pg/mL)	8.24 (6.47–10.73)	15.43 (11.73–21.61)	7.82 (6.37–10.27)	<0.0001	0.447	<0.0001 (OVC > Control)	1.0000 (ns)	<0.0001 (OCC < OVC)
VEGF (mU/mL)	700.77 (272.89–1,457.89)	1,819.46 (1,294.08–1,976.39)	1,775.15 (921.31–1,866.69)	<0.0001	0.162	<0.0001 (OVC > Control)	0.0022 (OCC > Control)	0.3029 (ns)
Amino acids and nitrogen metabolism								
α-AAs (mmol/L)	3.76 (3.65–4.01)	4.24 (3.92–5.11)	3.86 (3.72–4.14)	<0.0001	0.204	<0.0001	0.1075 (ns)	0.0012 (OCC < OVC)
ICs (mmol/L)	0.339 (0.197–0.518)	0.094 (0.035–0.190)	0.239 (0.135–0.396)	<0.0001	0.192	<0.0001 (OVC < Control)	0.2003 (ns)	0.0007 (OCC > OVC)
Arg (nmol/L)	44.09 (20.42–104.02)	21.52 (12.12–32.36)	25.74 (16.71–51.36)	0.0019	0.090	<0.0001 (OVC < Control)	0.0665 (ns)	0.3156 (ns)
Pro (nmol/L)	250.09 (106.78–521.30)	93.48 (48.89–209.11)	73.44 (38.08–176.59)	0.0004	0.104	0.0002 (OVC < Control)	<0.0001 (OCC < Control)	1.0000 (ns)

(*Cont'd...*)

Table 2. (*Continued*)

Predictor	Control (*n* = 75)	Ovarian cancer (*n* = 89)	Oral cavity cancer (*n* = 49)	Global Kruskal–Wallis test		*p*-value (Dunn post hoc comparisons; Bonferroni-adjusted)		
				p-value (Bonferroni-adjusted)	ε^2	Control vs. ovarian cancer (OVC)	Control vs. oral cavity cancer (OCC)	OVC vs. OCC
Gln (nmol/L)	95.96 (55.47–172.88)	60.10 (21.62–108.36)	28.51 (14.58–75.70)	0.0123	0.081	0.0149 (OVC < Control)	0.0002 (OCC < Control)	0.3003 (ns)
Urea (mmol/L)	6.85 (5.28–9.21); *n* = 72	7.13 (4.80–8.72); *n* = 86	4.09 (2.76–5.16); *n* = 48	<0.0001	0.122	1.0000 (ns)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
NH_4^+ (mg/L)	105.40 (70.99–168.06)	139.00 (91.31–209.80)	100.81 (70.98–177.55)	1.0000	0.026	0.1230 (ns)	1.0000 (ns)	0.0736 (ns)
Tumor-associated markers								
CA27.29 (mU/mL)	1.27 (0.69–2.75)	5.77 (0.74–10.80)	4.07 (2.04–8.13)	0.0120	0.080	0.0006 (OVC > Control)	0.0023 (OCC > Control)	1.0000 (ns)
CA15-3 (mU/mL)	28.46 (16.70–38.39)	99.92 (28.61–168.96)	26.89 (12.28–38.74)	<0.0001	0.183	<0.0001 (OVC > Control)	1.0000 (ns)	<0.0001 (OCC < OVC)
MCA (mU/mL)	12.49 (7.92–20.26)	57.97 (23.62–74.62)	12.81 (4.18–38.55)	<0.0001	0.269	<0.0001 (OVC > Control)	1.0000 (ns)	<0.0001 (OCC < OVC)
Salivary inorganic ion composition								
Na^+ (mg/L)	128.25 (93.08–163.23)	92.92 (67.40–146.65)	98.34 (60.83–126.90)	0.2343	0.045	0.0243 (OVC < Control)	0.0124 (OCC < Control)	1.0000 (ns)
Ca^{2+} (mg/L)	46.40 (30.89–60.63)	52.57 (34.45–75.98)	23.33 (14.00–33.14)	<0.0001	0.258	0.5905 (ns)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
F^- (mg/L)	3.06 (1.44–5.09)	1.48 (0.61–2.98)	0.62 (0.38–0.79)	<0.0001	0.253	0.0038 (OVC < Control)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
PO_4^{3-} (mg/L)	272.95 (225.80–340.58)	273.75 (164.13–404.88)	168.10 (105.80–232.55)	<0.0001	0.141	0.7529 (ns)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
Lipid peroxidation								
DC (c.u.)	2.74 (2.63–3.01)	2.84 (2.52–3.10)	2.42 (2.31–2.55)	<0.0001	0.194	1.0000 (ns)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
FTIR-derived lipid features								
1396H (a.u.)	0.79 (0.42–1.57)	1.07 (0.81–1.77)	0.58 (0.38–0.76)	0.0001	0.148	0.0380 (OVC > Control)	0.0198 (OCC < Control)	<0.0001 (OCC < OVC)
1458H (a.u.)	2.75 (1.92–4.01)	3.05 (2.24–4.51)	1.64 (1.35–2.04)	<0.0001	0.267	0.7865 (ns)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
1458S (a.u.)	12.00 (7.67–18.80)	13.10 (9.82–20.45)	6.84 (5.57–8.73)	<0.0001	0.248	0.8195 (ns)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
2853H (a.u.)	6.89 (6.17–9.51)	11.50 (8.78–14.30)	3.51 (3.21–4.15)	<0.0001	0.633	0.0115 (OVC > Control)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)

(*Cont'd...*)

Table 2. (*Continued*)

Predictor	Control (*n* = 75)	Ovarian cancer (*n* = 89)	Oral cavity cancer (*n* = 49)	Global Kruskal–Wallis test		*p*-value (Dunn post hoc comparisons; Bonferroni-adjusted)		
				p-value (Bonferroni-adjusted)	ε^2	Control vs. ovarian cancer (OVC)	Control vs. oral cavity cancer (OCC)	OVC vs. OCC
2853S (a.u.)	127.0 (111.0–188.5)	224.0 (154.0–280.0)	59.2 (55.6–70.9)	<0.0001	0.627	0.0327 (OVC > Control)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
2923H (a.u.)	11.70 (10.40–16.50)	18.60 (14.10–23.30)	5.74 (5.13–6.79)	<0.0001	0.630	0.0236 (OVC > Control)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
2923S (a.u.)	299.0 (264.5–446.5)	474.0 (354.0–594.0)	135.0 (117.0–160.0)	<0.0001	0.613	0.0673 (ns)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
2957S (a.u.)	20.10 (16.30–59.85)	33.20 (24.70–50.50)	10.20 (6.34–15.70)	<0.0001	0.462	0.2478 (ns)	<0.0001 (OCC < Control)	<0.0001 (OCC < OVC)
2923/2957H (a.u.)	4.46 (3.79–5.19)	5.50 (4.39–6.00)	5.34 (4.39–6.61)	0.0464	0.073	0.0116 (OVC > Control)	0.0031 (OCC > Control)	1.0000 (ns)

Notes: Data are presented as medians and interquartile ranges, Me (Q1–Q3). Overall differences between the three groups were assessed using the Kruskal–Wallis test with ε^2 as the effect-size measure, followed by Dunn's pairwise comparisons with Bonferroni adjustment; the actual number of available observations is reported for variables containing missing values. The *p*-value should be interpreted descriptively along with the results of multivariate maximum likelihood. MCP-1 values were missing because the reagent ran out, and the study was performed on only four plates (ovarian cancer: 89 samples in duplicate on two plates; oral cancer and control: 44/45 = 89 samples in duplicate on two more plates). For urea, missing values were one for oral cancer, three for ovarian cancer, and three for the control, due to insufficient sample size, as this study was performed last.
Abbreviations: AA: Amino acid; a.u.: Absorption units; DC: Diene conjugate; IC: Imidazole compound; IFN: Interferon; IL: Interleukin; MCA: Mucin-like carcinoma-associated antigen; TNF-α: tumor necrosis factor alpha; VEGF: Vascular endothelial growth factor.

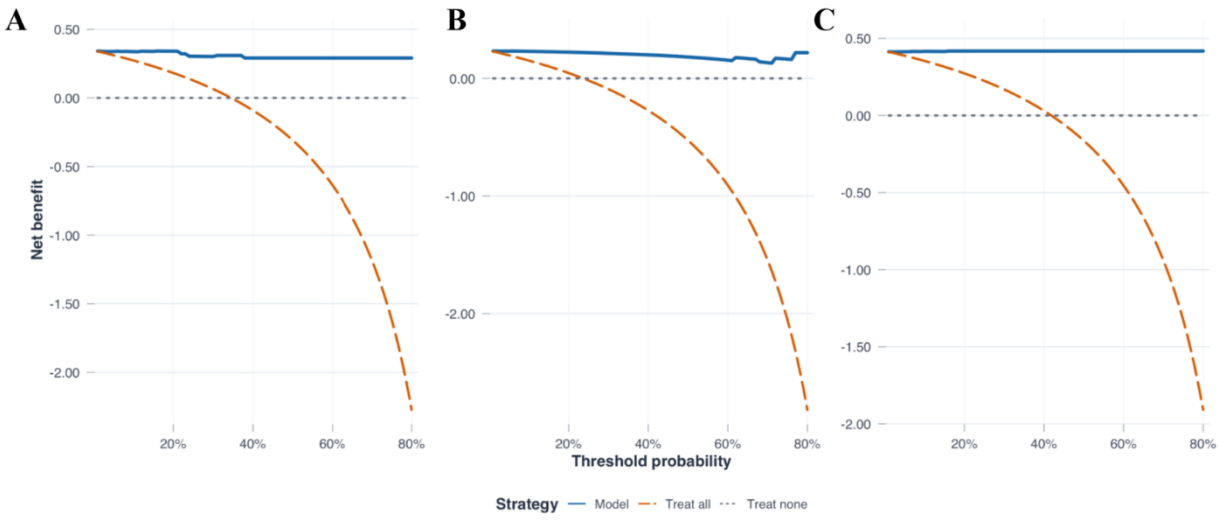


Figure 5. Decision curve analysis of class-specific extreme gradient boosting (XGBoost) predictions performed using a one-versus-rest approach for each diagnostic class. (A) Control versus all other classes. (B) Oral cavity cancer versus all other classes. (C) Ovarian cancer versus all other classes. Across most of the evaluated threshold-probability range, the XGBoost model provided greater net benefit than the treat-all and treat-none reference strategies, supporting its potential clinical utility while requiring confirmation through external validation.

net benefit across a wide range of threshold probabilities. This is particularly important because the analysis was conducted after removing the suspicious dominant feature 2957H. Therefore, even without this feature, the model retains potential clinical value in classifying oral cancer. For the control and ovarian cancer classes, the model also consistently outperforms alternative strategies across the entire presented threshold range. Especially for ovarian cancer, the model line remains quite flat and positive, indicating a consistent benefit of the classifier. Thus, DCA confirms that the model demonstrates good statistical metrics and has the potential to be useful from a clinical perspective. The results are preliminary, and external validation in an independent sample is required to confirm practical applicability.

4. Discussion

4.1. Biochemical profile of saliva in the control group

The control group had less pronounced angiogenic-immunoregulatory activity and lower total free amino acid content, against a background of relatively high arginine, glutamine, proline, and imidazole compounds. In addition, the saliva of healthy subjects had a more stable nitrogen and mineral buffer profile, as well as higher content and better preservation of long-chain lipid structures. Thus, low values of total α-amino acids mainly supported the classification of samples as controls, while high values reduced the model yield for the control class. A potential source of the increase in free amino acids may be increased proteolysis of salivary and tissue proteins, cell damage and renewal, and metabolic remodeling, but the direct mechanism was not determined in the present study.[41-43] Elevated values of arginine, glutamine, and histidine also supported the classification of the samples as controls. Arginine is known to support two metabolic pathways. Arginine can serve as a substrate for the formation of citrulline and nitric oxide by nitric oxide synthase, or for the formation of ornithine via arginase-1 for the subsequent formation of polyamines. Thus, relatively low arginine values may be associated with the restructuring of arginine metabolism. The proline content in saliva represents an integral characteristic of amino acid metabolism associated with the metabolism of glutamate, glutamine, arginine, and ornithine; the renewal of the collagen-containing extracellular matrix; and energy and redox regulation.[44-47] In oral fluid, its level may also depend on the secretory activity of the salivary glands and the metabolism of proline-rich proteins.[48] Glutamine can simultaneously reflect nitrogen consumption for proliferation, the intensity of glutaminolysis, nucleotide synthesis, antioxidant protection, and general protein catabolism.[49] It has been shown that the imidazole alkaloids epiisopylothurin and epiisopylosin suppress the formation of reactive oxygen species and degranulation of activated neutrophils, reduce the production of IL-6, and, in an experimental model of inflammation *in vivo*, reduce the level of myeloperoxidase.[50] It is known that imidazole compounds are capable of inhibiting the nuclear factor-kappa B (NF-κB) signaling pathway. This leads to a blockade of the nuclear translocation of p65 subunits, thereby decreasing the expression of proinflammatory cytokines, such as TNF-α, IL-1β, and IL-6.[51] Histidine-containing imidazole compounds, primarily carnosine, are capable of binding reactive aldehyde products of lipid peroxidation, including 4-hydroxynonenal, acrolein, and malondialdehyde, chelating transition metal ions, and limiting the formation of carbonyl adducts and protein cross-links.[52-55] The control profile was characterized by a relatively small free pool of α-amino acids with higher urea content in saliva. This combination may reflect efficient uptake and utilization of free amino acids followed by the utilization of their amino nitrogen in the form of urea. High values of inorganic phosphate predominantly increased model yield for the control class, whereas low values decreased it. A similar, but less pronounced and non-monotonic, trend was observed for fluoride. The resulting profile may reflect a more preserved mineral-buffering and remineralizing state of saliva in individuals in the control group, since phosphate helps regulate acid-base balance and, together with calcium, maintains saliva saturation relative to enamel minerals, and fluoride reduces demineralization and promotes remineralization.[56,57] A decrease in salivary fluoride concentration in cancer may reflect a decrease in the local availability of fluoride ions, which are involved in suppressing demineralization and maintaining remineralization of hard dental tissues.[58] In oral cancer, such a change may be associated with local restructuring of the secretion and mineral composition of saliva, the state of the mucous membrane, changes in the microbiota, and oral hygiene.[59,60] In ovarian cancer, a decrease in salivary fluoride is most likely a nonspecific consequence of systemic, secretory, behavioral, or therapeutic factors, since a direct mechanism for the influence of ovarian tumor on fluoride metabolism in saliva has not been established.[26]

The IL-6 protein normally coordinates inflammatory and acute-phase responses, immune defense, tissue regeneration, and metabolic adaptation.[61-63] During chronic tumor activation, the IL-6/Janus kinase (JAK)/signal transducer and activator of transcription 3 (STAT3) pathway can support tumor cell proliferation, survival, angiogenesis, myeloid cell recruitment, and metastasis.[64,65] The effect of salivary IL-6 on control is nonlinear in the graph, so there is no simple direction in which low

IL-6 corresponds to normal. IL-10 limits excessive inflammation and protects tissues from immune damage. In tumors, its role is dual: IL-10 can suppress antitumor immunity by enhancing immunosuppressive mechanisms, but under some conditions it can activate the metabolism and function of exhausted $CD8^+$ T cells.[66] In the obtained model, low salivary IL-10 predominantly supports the control class, which may correspond to the absence of pronounced tumor-associated immunoregulatory activation. IL-18 belongs to the IL-1 family of cytokines and is involved in mucosal immune surveillance, linking inflammasome recognition of microbial and damaging signals to natural killer (NK) and T cells, IFN-γ production, and antimicrobial defense.[67-69] At the same time, epithelial IL-18 contributes to barrier homeostasis and host–microbiota interactions.[70] Therefore, IL-18 content in saliva can be considered an integral characteristic of innate immune recognition, epithelial–immune interactions, and the readiness of the local cell-mediated response, but not an unambiguous indicator of inflammation intensity. VEGF stimulates the proliferation and migration of endothelial cells, supports physiological angiogenesis and tissue repair, and regulates vascular permeability. In oncology, its increase, often associated with hypoxia, promotes tumor vessel formation, oxygen and nutrient supply to the tumor, and tumor progression.[71,72] In Figure 4, low salivary VEGF maintains control, which may reflect less pronounced angiogenic activation. The control group was defined by a contribution of 2,853 cm^{-1} in saliva. This can be interpreted as the fact that in the control saliva, intact long fatty acid chains are relatively better preserved; the total contribution of normal aliphatic lipid structures (phospholipids of cell membranes, sphingolipids, triacylglycerols, free fatty acids, cholesterol esters, lipoprotein complexes, and extracellular vesicles) is higher; there is less oxidative fragmentation of lipids; and the normal composition of salivary lipids and lipoprotein complexes is preserved.

Normally, phosphate ions in saliva are involved in pH regulation and maintain the conditions necessary for calcium–phosphate balance on the tooth surface.[73] This allows oral fluid to serve as a source of mineral components that can restore areas of initial enamel demineralization.[74] Fluoride exerts a protective effect primarily in the oral cavity.[75] It reduces the solubility of the mineral phase in an acidic environment and promotes the subsequent precipitation of calcium and phosphate.[76] Therefore, the combined content of phosphate and F^- reflects the state of mineralizing and anticaries protection of saliva.

CA27.29 is an immunoreactive epitope of the extracellular region of the transmembrane mucin (MUC1).[77] This indicator depends not only on the amount of MUC1 but also on the degree of its glycosylation and the availability of the recognized epitope.[78] Its physiological presence in saliva may be associated with MUC1 expression by epithelial cells of the oral cavity and salivary gland ducts, renewal of the epithelial glycocalyx, and release of soluble mucin fragments.[79,80] Since MUC1 is involved in the formation of the protective mucosal surface, retention of secreted mucins, and maintenance of the barrier properties of the epithelium, the content of CA27.29 may indirectly characterize the state of the mucin component of the salivary–epithelial barrier.[81-83]

4.2. Salivary biochemical profile in ovarian cancer patients

The group of patients diagnosed with ovarian cancer had higher levels of TNF-α, MCP-1, CA15-3, and IL-8, which consistently increased the model yield for this group. A less pronounced positive effect was observed for IFN-α and IFN-γ. Additionally, low levels of imidazole compounds weakly supported the classification as ovarian cancer. IL-6 was the most significant predictor, although it demonstrated a pronounced nonlinear relationship. TNF-α is responsible for acute inflammation and the activation of endothelium and immune cells. Through tumor necrosis factor receptor 1 (TNFR1)/TNFR2, it can stimulate cell survival or death. When chronically activated, it maintains tumor-associated inflammation, NF-κB signaling, angiogenesis, and tissue damage. At high local concentrations, it can be cytotoxic. In distant tumors, it mainly affects the systemic inflammatory reaction and the response of the salivary glands.[84] MCP-1 (C-C motif chemokine ligand 2 [CCL2]) attracts monocytes, some T cells, and other immune cells to the site of injury. It can promote monocyte recruitment and the formation of tumor-associated macrophages, including an M2-like immunosuppressive phenotype that supports invasion and angiogenesis. It also promotes mucosal monocyte-macrophage activation.[85-87] CA15-3 reflects fragments of MUC1, a membrane mucin involved in epithelial cell protection and signaling. During tumor transformation, MUC1 can be overexpressed, abnormally glycosylated, and more actively shed from the cell surface.[88] Elevated salivary CA15-3 in ovarian cancer may reflect systemically induced overexpression, abnormal glycosylation, and increased release of MUC1, as well as altered salivary gland secretory function induced by circulating inflammatory mediators.[89] Elevated MUC1-associated markers may be associated with impaired normal apical secretion and reduced protective properties of the secretory mucins MUC5B and MUC7. This mechanism is biologically plausible but has not yet been proven directly in ovarian cancer and requires parallel measurement of MUC1, MUC5B, MUC7, salivary

flow rate, and serum CA15-3.[90] IL-8 attracts and activates neutrophils, providing early protection against infection and participating in inflammation; it may also support neutrophilic inflammation, angiogenesis, invasion, and migration of tumor cells. In the ovarian tumor environment, IL-8 can also be produced by immune and stromal cells.[91] IFN-α, or type I IFN, creates an antiviral state, activates NK cells and innate immune mechanisms, and limits cell proliferation. It can support antitumor immune activation and directly inhibit proliferation, but long-term IFN signaling in some tumors is associated with adaptation and immune resistance.[92,93] IFN-γ, or type II IFN, is produced primarily by T and NK cells and activates macrophages, enhances antigen presentation, and supports antitumor immune surveillance. It may reflect an active Th1/NK response and cytotoxic immunity. At the same time, chronic IFN-γ is capable of inducing programmed death ligand 1 (PD-L1) and adaptive tumor evasion of the immune response.[94,95] In ovarian cancer, IL-6 acquires a different biological meaning than in the control group, where the biological significance of the cytokine was discussed. This difference is because long-term activation of the IL-6/JAK/STAT3 signaling pathway can support the survival and proliferation of tumor cells, angiogenesis, invasion, and the development of drug resistance. Tumor-associated fibroblasts may be a source of IL-6 in the tumor microenvironment, indicating the involvement of the stromal component in maintaining chronic cytokine signaling.[96,97] However, IL-6 levels in saliva cannot be directly equated with its production in tumor tissue, since it also depends on the systemic inflammatory response, the state of the oral mucosa, and the function of the salivary glands.[98,99] Thus, changes in TNF-α, MCP-1, IL-8, and IL-6 in saliva primarily characterize inflammatory activation and myeloid cell recruitment. IFN-α and IFN-γ characterize the state of the innate, Th1-, and NK-associated immune response. CA15-3 characterizes the intensity of expression and release of MUC1-containing epithelial material. In ovarian cancer, these represent a systemically mediated immune-inflammatory and mucin-associated salivary signature.

In patients with ovarian cancer, saliva IL-10 levels increased significantly, whereas only numerical trends toward increases were observed for IL-2 and IL-4. IL-2 is involved in antigen-dependent activation and proliferation of effector T lymphocytes and is also necessary for maintaining regulatory T cells.[100] IL-2 is synthesized mainly by activated T lymphocytes after antigen recognition. Binding of IL-2 to the IL-2 receptor triggers the JAK1/JAK3–STAT5, phosphoinositide 3-kinase (PI3K)–AKT, and RAS–extracellular signal-regulated kinase (ERK) cascades, which regulate the cell cycle, survival, and proliferation of lymphocytes. Therefore, IL-2 primarily reflects the activation and clonal expansion of T cells.[101,102] Normally, IL-2 supports the proliferation of activated $CD4^+$- and $CD8^+$-lymphocytes and the development of a cytotoxic response. Moreover, it is necessary for the maintenance of forkhead box P3 $(FOXP3)^+$ regulatory T cells, which limit autoimmune and excessive inflammatory reactions.[103,104] IL-4 is one of the main regulators of the type II immune response. Through the IL-4Rα receptor and STAT6 activation, it induces the transcription factor GATA binding protein 3 (GATA3) and promotes the differentiation of $CD4^+$ lymphocytes in the Th2 direction.[105] IL-4 also affects B lymphocytes. Together with the CD40 signal, it promotes immunoglobulin class switching, particularly the formation of immunoglobulin E and immunoglobulin G4. Therefore, IL-4 is associated with humoral defense, the antiparasitic response, and allergic reactions.[106,107] In macrophages, IL-4 activates the STAT6-dependent alternative activation program. This program is associated with limiting acute inflammatory reactions, tissue repair, extracellular matrix remodeling, and barrier restoration. Current data indicate that this is not a single, strictly fixed "M2" state, but a spectrum of context-dependent macrophage phenotypes.[108,109] IL-4 is associated with Th2 differentiation, humoral response, and alternative macrophage activation programs, while IL-10 limits the production of proinflammatory mediators and the activity of antigen-presenting cells.[110] IL-10 is a central cytokine that limits excessive inflammation. Through the IL-10 receptor and STAT3, it suppresses the production of proinflammatory cytokines and chemokines by macrophages and other antigen-presenting cells, reduces the intensity of antigen presentation, and limits excessive T-cell activation.[111] IL-10 also promotes metabolic restructuring in macrophages. Experimental studies have shown that IL-10 prevents the inflammatory transition to increased glycolysis, suppresses mechanistic target of rapamycin (mTOR)-dependent processes, supports the removal of damaged mitochondria, and thereby limits reactive oxygen species formation and NLR family pyrin domain containing 3 (NLRP3) inflammasome activation.[112] A predominant increase in IL-10 potentially indicates enhanced mechanisms that limit chronic inflammation but, in the context of tumor progression, can also weaken effective antitumor immune surveillance.[113,114] An increase in salivary VEGF in ovarian cancer is consistent with the simultaneous activation of the inflammatory–angiogenic cytokine network. IL-6, IL-8, TNF-α, and MCP-1 support the expression of proangiogenic factors, the recruitment of myeloid cells, stromal remodeling, and increased vascular permeability.[115-117] In this system, VEGF may act as an effector component linking chronic inflammation to

the formation of a pathological vasculature.[118] Therefore, the identified profile characterizes not unidirectional inflammation, but a dynamic balance between tumor-supporting angiogenic signaling and the attempt of the immune system to limit its consequences.

In patients with ovarian cancer, increased MCA content in saliva may reflect increased formation and release of mucin-like epithelial glycoproteins, as well as changes in their *O*-glycosylation and the availability of recognizable epitopes.[119-121] An increase in the concentration of NH_4^+ may be associated with intensified nitrogen catabolism and bacterial breakdown of urea and arginine, which regulate the acid–base balance of the oral cavity.[122,123] An increase in the total content of imidazole compounds may indicate a restructuring of the metabolism of histidine and its derivatives associated with buffering, antioxidant, and immunoregulatory processes.[124,125] Together, these changes characterize a multicomponent metabolic signature, including mucin, nitrogen, and histidine-imidazole metabolism; however, because salivary parameters are nonspecific and imidazole compounds are determined analytically, we cannot establish their direct tumor origin.

4.3. Salivary biochemical profile in oral cancer patients

The SHAP analysis revealed that the classification of saliva samples as oral cancer–related was determined primarily by the low values of the 2,923 cm^{-1} band height and the 2,853 cm^{-1} band height and area. A similar, weaker trend was observed for the 1,458 cm^{-1} band. The 2,923 and 2,853 cm^{-1} bands predominantly correspond to asymmetric and symmetric stretching vibrations of CH_2 groups, while the 1,458 cm^{-1} region includes CH_2 bending vibrations in extended aliphatic regions of fatty acid chains. Therefore, the signature identified by the model indicates a reduced spectral contribution of salivary CH_2-containing lipid structures in oral cancer patients. These structures are part of phospholipids, sphingolipids, triacylglycerols, free fatty acids, lipoprotein complexes, and extracellular vesicle membranes. In a study of salivary exosomes from oral cancer patients, differences between the tumor and control groups were also found in the regions of 2,924 and 2,854 cm^{-1}, which were attributed to membrane lipids.[126] In that study, these bands were less intense in exosomes from oral cancer patients than in healthy individuals, consistent with the spectral signature we identified. An additional component of the lipid signature was a decrease in the height of the 1,396 cm^{-1} band compared with the control group. This region is predominantly associated with symmetric deformation vibrations of CH_3 groups, although in complex biological samples vibrations of carboxylate groups and amino acid side chains may also contribute. Therefore, the decrease in the height of the 1,396 cm^{-1} band may reflect a reduced relative contribution of terminal and side methyl groups in lipid and lipoprotein structures or a change in their molecular environment. A nearby region around 1,404 cm^{-1} was previously attributed in salivary exosome spectra to C–H deformation vibrations in acyl residues of lipids and amines, confirming the biological information content of this spectral range but also indicating its chemical nonspecificity.[127]

At the same time, patients with oral cancer showed a higher 2,923/2,957 band height ratio than the control group. Since the 2,923 cm^{-1} band predominantly reflects asymmetric vibrations of the CH_2 groups, and the 2,957 cm^{-1} band reflects asymmetric vibrations of the CH_3 groups, an increase in this ratio indicates a relative shift in the spectral balance toward the methylene component. In simplified lipid systems, a higher CH_2/CH_3 ratio may correspond to a higher proportion of internal methylene units relative to terminal methyl groups and may be associated with changes in the length or degree of branching of aliphatic chains.[128,129] However, given the simultaneous decrease in the absolute intensity of the 2,923 and 2,853 cm^{-1} bands, the result obtained should not be interpreted as an increase in the total content of CH_2-containing lipids. It is more likely that oral cancer shows a decrease in both the methylene and methyl components, but the contribution of CH_3-linked structures decreases relatively more, leading to an increase in the 2,923/2,957 ratio. Relative intensities of lipid bands have already been used to identify structural differences in salivary exosome membranes in oral cancer.[130,131] Diene conjugates are predominantly early products of polyunsaturated fatty acid peroxidation, arising from the rearrangement of double bonds in the initial stages of free radical damage to lipids.[132] The reduced content of diene conjugates may be associated not with the absence of oxidative stress, but with a decrease in available polyunsaturated lipid substrate or its rapid conversion to later oxidation products.[133] Together with the decrease in the intensity of the bands at 2,923, 2,853, and 1,458 cm^{-1}, this may indicate quantitative and qualitative remodeling of the lipid component of saliva in oral cancer. Thus, the combination of decreased values of 2923H, 2853H/S, 1458H, and 1396H with an increased 2923/2957H ratio may reflect not a simple quantitative decrease in lipids, but a qualitative remodeling of the aliphatic component of saliva: a change in the ratio of CH_2- and CH_3-containing structures, the composition of membrane lipids, lipoprotein complexes, and extracellular vesicles. This may be associated with several complementary processes.[134] First, tumor transformation is accompanied by restructuring the synthesis, absorption, and utilization of fatty acids necessary for proliferation, formation of

new membranes, and intracellular signaling. This can lead to a redistribution of lipids among tumor tissue, salivary secretion, cellular sediment, and extracellular vesicles. Second, rapid renewal of the tumor epithelium, apoptosis, necrosis, and inflammatory infiltration alter the amount and composition of cellular membrane material that directly enters saliva. Third, the secretory function of the salivary glands and the composition of the lipid–protein complexes they secrete can change.[135,136] Salivary extracellular vesicles have a lipid bilayer and carry lipids, proteins, and nucleic acids.[137] An additional mechanism may be oxidative damage to fatty acid chains. Reactive oxygen species initiate peroxidation of polyunsaturated phospholipids to form hydroperoxides, malondialdehyde, and other reactive carbonyl products.[138] The cleavage and fragmentation of acyl chains can reduce the number of extended $–CH_2–CH_2–$ sequences, changing the intensity of the 2,923, 2,853, and 1,458 cm^{-1} bands. It is more accurate to consider oxidative fragmentation as one of the potential components of general lipid remodeling. Changes in fatty acid chain composition can affect membrane fluidity, permeability, and organization, as well as receptor localization, intercellular signaling, and extracellular vesicle formation.[139]

Interferon-gamma is produced primarily by Th1, NK, and activated $CD8^+$ cells and enhances antigen presentation, cytotoxic responses, and classical macrophage activation. Therefore, its decrease in saliva may indicate a weakened local Th1 cytotoxic signature. In studies of oral squamous cell carcinoma, downregulation of IFN-γ and other Th1-associated factors has been described in the formation of an immunosuppressive tumor microenvironment.[140-142] MCP-1 is a chemokine that recruits C-C motif chemokine receptor 2 $(CCR2)^+$ monocytes, which, after migrating into tumor tissue, can differentiate into tumor-associated macrophages. Therefore, a decrease in MCP-1 in saliva may indicate a decrease in the soluble monocyte-recruiting signal specifically in oral fluid. However, this does not prove a decrease in macrophage numbers in the tumor. In the tissue microenvironment, MCP-1 can be produced locally by tumor and stromal cells and consumed to form a chemotactic gradient, and saliva concentration does not necessarily reflect tissue expression.[143-145] IL-4 is a Th2-associated cytokine that promotes alternative, M2-like macrophage activation associated with inflammation limitation, tissue repair, angiogenesis, and immunosuppressive properties.[146,147] Therefore, a combination of low IFN-γ with a tendency toward increased IL-4 can be viewed as a relative shift of the Th1/Th2 and M1/M2 axis toward a Th2/M2-like response, potentially favoring immune evasion and remodeling of the tumor microenvironment. In oral cancer, a similar combination of IFN-γ suppression and increased IL-4-associated expression has previously been observed, predominantly with higher tumor burden and lymph node involvement.[148,149] Thus, the combined decrease in MCP-1 and IFN-γ, with a slight tendency toward an increase in IL-4, may reflect a restructuring of the local immune response toward less pronounced cell-mediated antitumor activity.

A decrease in Ca^{2+} and F^- levels may reflect a change in the inorganic composition of salivary secretions and a decrease in the delivery of mineral components to the oral fluid. For Ca^{2+}, this may be associated with changes in acinar secretion, ductal modification of primary saliva, protein binding, or the relative contribution of individual salivary glands.[150-152] A decrease in F^- is a less specific sign of secretory dysfunction, since its content in saliva is largely determined by external intake and recent use of fluoride-containing agents.[76]

Limitations of this study include the small sample size and the predominance of advanced cancer stages, particularly in oral cancer. The distribution of patients by disease stage may significantly affect both the estimated diagnostic performance of the model and its potential for future use in early cancer detection. Theoretically, more advanced tumor progression may be accompanied by more pronounced local and systemic metabolic, inflammatory, and biochemical changes, which could increase the discriminability between the cancer and control groups. In this case, model performance calculated on a sample with a predominance of advanced stages could potentially overestimate its performance in patients with early-stage disease. To test this hypothesis, the most informative approach would be a stage-specific sensitivity analysis, such as evaluating the model separately for early (I–II) and advanced (III–IV) stages or training and testing the model in the corresponding strata. Such an analysis would allow us to determine whether the diagnostic signal is maintained at early stages or whether part of the model's discriminatory ability reflects the severity of changes in advanced tumor progression. However, in the current sample, such an analysis does not appear statistically feasible. After stratification by stage, the number of observations in individual subgroups, particularly in the oral cancer group, becomes too small to reliably train and evaluate a multivariable model. This would lead to high variability in estimates, wide confidence intervals, instability in sensitivity and specificity, and an increased risk of overfitting.

5. Conclusion

The XGBoost model correctly classified all cases of ovarian

and oral cancer. A permutation test showed that the original model was significantly superior to models with randomly shuffled labels. SHAP analysis revealed that the model's recognition of the control group was associated with features reflecting the relative preservation of salivary metabolic homeostasis: amino acid metabolism, redox and acid–base balance, moderate immunoinflammatory activity, the functional state of buffer systems, and aliphatic lipid and lipoprotein components. For the ovarian cancer class, cytokine and tumor-associated indicators contributed most to prognosis, suggesting a systemically mediated immunoinflammatory signature in saliva. Oral cancer classification was determined primarily by the spectral characteristics of the lipid component, potentially reflecting fatty acid chain remodeling, changes in membranogenesis, the composition of cellular and vesicular material released into saliva, and possible oxidative lipid modification. It should be emphasized that SHAP characterizes the contribution of individual features to model decisions but does not prove causal biological mechanisms or the direction of intergroup differences without comparison with the underlying data.

Although simultaneously discriminating between ovarian and oral cancer is not a typical clinical task, the one-versus-all DCA results demonstrate that the model retains positive clinical value for each diagnostic category, highlighting the potential applicability of the approach to more realistic binary screening and triage tasks. Furthermore, the multiclass model has independent research validity, as it allows simultaneous comparison of different salivary signatures: for ovarian cancer, the leading predictors were predominantly immunoinflammatory markers (IL-6, TNF-α, MCP-1), whereas for oral cancer, it was predominantly FTIR spectral signatures. Thus, the multiclass model was used primarily to assess the differences between systemic and local components of salivary profiles, and its clinical application in specific binary scenarios requires separate validation.

Decision curve analysis revealed that the XGBoost model demonstrates positive net clinical benefit for all three classes across a wide range of threshold probabilities and significantly outperforms the "treat-all" and "treat-none" strategies. These results indicate the potential applicability of the model in supporting diagnostic decisions. At this stage, the results should be considered preliminary, as they represent internal validation estimates based on a single-center cohort and await external verification.

Acknowledgments

None.

Funding

This research was funded by the Russian Science Foundation (grant number 23-15-00188-П).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: All authors
Formal analysis: Lyudmila V. Bel'skaya
Investigation: All authors
Methodology: Elena I. Dyachenko
Writing – original draft: Elena I. Dyachenko
Writing – review & editing: Lyudmila V. Bel'skaya

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Omsk State Pedagogical University (20 March 2024; protocol code 46-04/2). Written informed consent was obtained from the patient included in the study.

Consent for publication

The patient provided written informed consent for the publication of anonymized data and any accompanying images.

Availability of data

The original contributions presented in this study are included in the article. Further inquiries can be directed to the author.

References

1. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov*. 2022;12(1):31-46.

   doi: 10.1158/2159-8290.CD-21-1059

2. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun*. 2020;11(1):5120.

   doi: 10.1038/s41467-020-18794-x

3. Henry NL, Hayes DF. Cancer biomarkers. *Mol Oncol*. 2012;6(2):140-146.

   doi: 10.1016/j.molonc.2012.01.010

4. Pepe MS, Etzioni R, Feng Z, *et al*. Phases of biomarker development for early detection of cancer. *J Natl Cancer Inst*. 2001;93(14):1054-1061.

   doi: 10.1093/jnci/93.14.1054

5. Diamandis EP. Cancer biomarkers: can we turn recent failures into success? *J Natl Cancer Inst*. 2010;102(19):1462-1467.

doi: 10.1093/jnci/djq306

6. Yurkovetsky Z, Skates S, Lomakin A, *et al*. Development of a multimarker assay for early detection of ovarian cancer. *J Clin Oncol*. 2010;28(13):2159-2166.

doi: 10.1200/JCO.2008.19.2484

7. Wang Q, Gao P, Wang X, Duan Y. Investigation and identification of potential biomarkers in human saliva for the early diagnosis of oral squamous cell carcinoma. *Clin Chim Acta*. 2014;427:79-85.

doi: 10.1016/j.cca.2013.10.004

8. Jain A, Kotimoole CN, Ghoshal S, *et al*. Identification of potential salivary biomarker panels for oral squamous cell carcinoma. *Sci Rep*. 2021;11:3365.

doi: 10.1038/s41598-021-82635-0

9. Bzdok D, Krzywinski M, Altman N. Machine learning: a primer. *Nat Methods*. 2017;14(12):1119-1120.

doi: 10.1038/nmeth.4526

10. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods*. 2018;15(4):233-234.

doi: 10.1038/nmeth.4642

11. Song X, Yang X, Narayanan R, *et al*. Oral squamous cell carcinoma diagnosed from saliva metabolic profiling. *Proc Natl Acad Sci USA*. 2020;117(28):16167-16173.

doi: 10.1073/pnas.2001395117

12. Zhang T, Pang A, Lyu J, *et al*. Application of Nonlinear Models Combined with Conventional Laboratory Indicators for the Diagnosis and Differential Diagnosis of Ovarian Cancer. *J Clin Med*. 2023;12(3):844.

doi: 10.3390/jcm12030844

13. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010;33(1):1-22.

doi: 10.18637/jss.v033.i01

14. Breiman L. Random forests. *Mach Learn*. 2001;45:5-32.

doi: 10.1023/A:1010933404324

15. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY: Association for Computing Machinery; 2016:785-794.

doi: 10.1145/2939672.2939785

16. Couronné R, Probst P, Boulesteix AL. Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinformatics*. 2018;19(1):270.

doi: 10.1186/s12859-018-2264-5

17. Subramanian J, Simon R. Overfitting in prediction models: is it a problem only in high dimensions? *Contemp Clin Trials*. 2013;36(2):636-641.

doi: 10.1016/j.cct.2013.06.011

18. Collins GS, Moons KGM, Dhiman P, *et al*. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378.

doi: 10.1136/bmj-2023-078378

19. Tseng YJ, Wang YC, Hsueh PC, Wu CC. Development and validation of machine learning-based risk prediction models of oral squamous cell carcinoma using salivary autoantibody biomarkers. *BMC Oral Health*. 2022;22(1):534.

doi: 10.1186/s12903-022-02607-2

20. Li Y, St John MAR, Zhou X, *et al*. Salivary transcriptome diagnostics for oral cancer detection. *Clin Cancer Res*. 2004;10(24):8442-8450.

doi: 10.1158/1078-0432.CCR-04-1167

21. Hu S, Li Y, Wang J, *et al*. Human saliva proteome and transcriptome. *J Dent Res*. 2006;85(12):1129-1133.

doi: 10.1177/154405910608501212

22. Hu S, Xie Y, Ramachandran P, *et al*. Large-scale identification of proteins in human salivary proteome by liquid chromatography/mass spectrometry and two-dimensional gel electrophoresis-mass spectrometry. *Proteomics*. 2005;5(6):1714-1728.

doi: 10.1002/pmic.200401037

23. Yoshizawa JM, Schafer CA, Schafer JJ, Farrell JJ, Paster BJ, Wong DT. Salivary biomarkers: toward future clinical and diagnostic utilities. *Clin Microbiol Rev*. 2013;26(4):781-791.

doi: 10.1128/CMR.00021-13

24. Dawes C, Wong DTW. Role of Saliva and Salivary Diagnostics in the Advancement of Oral Health. *J Dent Res*. 2019;98(2):133-141.

doi: 10.1177/0022034518816961

25. Cuevas-Córdoba B, Santiago-García J. Saliva: a fluid of study for OMICS. *OMICS*. 2014;18(2):87-97.

doi: 10.1089/omi.2013.0064

26. Li Y, Ou Y, Fan K, Liu G. Salivary diagnostics: opportunities and challenges. *Theranostics*. 2024;14(18):6969-6990.

doi: 10.7150/thno.100600

27. Villa A, Wolff A, Aframian D, *et al*. World Workshop on Oral Medicine VI: a systematic review of medication-induced salivary gland dysfunction: prevalence, diagnosis, and treatment. *Clin Oral Investig*. 2015;19(7):1563-1580.

doi: 10.1007/s00784-015-1488-2

28. Mortazavi H, Yousefi-Koma AA, Yousefi-Koma H. Extensive comparison of salivary collection, transportation, preparation, and storage methods: a systematic review. *BMC Oral Health*. 2024;24:168.

doi: 10.1186/s12903-024-03902-w

29. Riis JL, Ahmadi H, Hamilton KR, Hand T, Granger DA. Best practice recommendations for the measurement and interpretation of salivary proinflammatory cytokines in biobehavioral research. *Brain Behav Immun*. 2021;91:105-116.

doi: 10.1016/j.bbi.2020.09.009

30. Esperouz F, Ciavarella D, Santarelli A, *et al*. Saliva-based biomarkers in oral squamous cell carcinoma using OMICS technologies: a systematic review. *Oral*. 2024;4(3):293-302.

doi: 10.3390/oral4030024

31. Banavar G, Ogundijo O, Toma R, *et al*. The salivary metatranscriptome as an accurate diagnostic indicator of oral cancer. *NPJ Genom Med*. 2021;6:105.

doi: 10.1038/s41525-021-00257-x

32. Lau C, Kim Y, Chia D, *et al*. Role of pancreatic cancer-derived exosomes in salivary biomarker development. *J Biol Chem*. 2013;288(37):26888-26897.

doi: 10.1074/jbc.M113.452458

33. Cui L, Zheng J, Lu Y, *et al*. New frontiers in salivary extracellular vesicles: transforming diagnostics, monitoring, and therapeutics in oral and systemic diseases. *J Nanobiotechnology*. 2024;22:171.

doi: 10.1186/s12951-024-02443-2

34. Tajmul M, Parween F, Singh L, *et al*. Identification and validation of salivary proteomic signatures for non-invasive detection of ovarian cancer. *Int J Biol Macromol*. 2018;108:503-514.

doi: 10.1016/j.ijbiomac.2017.12.014

35. Yang J, Xiang C, Liu J. Clinical significance of combining salivary mRNAs and carcinoembryonic antigen for ovarian cancer detection. *Scand J Clin Lab Invest*. 2021;81(1):39-45.

doi: 10.1080/00365513.2020.1852478

36. Lee YH, Kim JH, Zhou H, Kim BW, Wong DTW. Salivary transcriptomic biomarkers for detection of ovarian cancer: for serous papillary adenocarcinoma. *J Mol Med*. 2012;90(4):427-434.

doi: 10.1007/s00109-011-0829-0

37. Sarf EA, Yunkind KS, Solomatin DV, Bel'skaya LV. Sex-specific salivary biochemical composition in a healthy population. *Int J Mol Sci*. 2026;27(5):2214.

doi: 10.3390/ijms27052214

38. Bel'skaya LV, Sarf EA, Solomatin DV, Kosenok VK. Analysis of the lipid profile of saliva in ovarian and endometrial cancer by IR Fourier spectroscopy. *Vib Spectrosc*. 2019;104:102944.

doi: 10.1016/j.vibspec.2019.102944

39. Solomatin DV, Sarf EA, Bel'skaya LV. Metabolic features of saliva before and after breast cancer surgery. *Metabolites*. 2025;15(11):693.

doi: 10.3390/metabo15110693

40. Bel'skaya LV, Kosenok VK, Massard G. Endogenous intoxication and saliva lipid peroxidation in patients with lung cancer. *Diagnostics (Basel)*. 2016;6(4):39.

doi: 10.3390/diagnostics6040039

41. Liu X, Ren B, Ren J, Gu M, You L, Zhao Y. The significant role of amino acid metabolic reprogramming in cancer. *Cell Commun Signal*. 2024;22(1):380.

doi: 10.1186/s12964-024-01760-1

42. Neves LX, Granato DC, Busso-Lopes AF, *et al*. Peptidomics-Driven Strategy Reveals Peptides and Predicted Proteases Associated With Oral Cancer Prognosis. *Mol Cell Proteomics*. 2021;20:100004.

doi: 10.1074/mcp.RA120.002227

43. Syrjänen SM, Alakuijala L, Alakuijala P, Markkanen SO, Markkanen H. Free amino acid levels in oral fluids of normal subjects and patients with periodontal disease. *Arch Oral Biol*. 1990;35(3):189-193.

doi: 10.1016/0003-9969(90)90054-e

44. Guillard J, Schwörer S. Metabolic control of collagen synthesis. *Matrix Biol*. 2024;133:43-56.

doi: 10.1016/j.matbio.2024.07.003

45. Phang JM. Proline Metabolism in Cell Regulation and Cancer Biology: Recent Advances and Hypotheses. *Antioxid Redox Signal*. 2019;30(4):635-649.

doi: 10.1089/ars.2017.7350

46. Cappelletti P, Tallarita E, Rabattoni V, Campomenosi P, Sacchi S, Pollegioni L. Proline oxidase controls proline, glutamate, and glutamine cellular concentrations in a U87 glioblastoma cell line. *PLoS One*. 2018;13(4):e0196283.

doi: 10.1371/journal.pone.0196283

47. Vettore LA, Westbrook RL, Tennant DA. Proline metabolism and redox; maintaining a balance in health and disease. *Amino Acids*. 2021;53(12):1779-1788.

doi: 10.1007/s00726-021-03051-2

48. Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J*. 2000;14(2):231-241.

doi: 10.1096/fasebj.14.2.231

49. Yang L, Venneti S, Nagrath D. Glutaminolysis: A Hallmark of Cancer Metabolism. *Annu Rev Biomed Eng*. 2017;19:163-

194.

doi: 10.1146/annurev-bioeng-071516-044546

50. Rocha TM, Machado NJ, de Sousa JAC, *et al*. Imidazole alkaloids inhibit the pro-inflammatory mechanisms of human neutrophil and exhibit anti-inflammatory properties in vivo. *J Pharm Pharmacol*. 2019;71(5):849-859.

doi: 10.1111/jphp.13068

51. Dos Santos Nascimento MVP, Munhoz ACM, De Campos Facchin BM, *et al*. New preclinical evidence of anti-inflammatory effect and safety of a substituted fluorophenyl imidazole. *Biomed Pharmacother*. 2019;111:1399-1407.

doi: 10.1016/j.biopha.2019.01.052

52. Liu Y, Xu G, Sayre LM. Carnosine inhibits (E)-4-hydroxy-2-nonenal-induced protein cross-linking: structural characterization of carnosine-HNE adducts. *Chem Res Toxicol*. 2003;16(12):1589-1597.

doi: 10.1021/tx034160a

53. Velez S, Nair NG, Reddy VP. Transition metal ion binding studies of carnosine and histidine: biologically relevant antioxidants. *Colloids Surf B Biointerfaces*. 2008;66(2):291-294.

doi: 10.1016/j.colsurfb.2008.06.012

54. Lavilla CJ Jr, Billacura MP, Hanna K, *et al*. Carnosine protects stimulus-secretion coupling through prevention of protein carbonyl adduction events in cells under metabolic stress. *Free Radic Biol Med*. 2021;175:65-79.

doi: 10.1016/j.freeradbiomed.2021.08.233

55. Berdaweel IA, Monroe TB, Alowaisi AA, *et al*. Iron scavenging and suppression of collagen cross-linking underlie antifibrotic effects of carnosine in the heart with obesity. *Front Pharmacol*. 2024;14:1275388.

doi: 10.3389/fphar.2023.1275388

56. Enax J, Fandrich P, Schulze zur Wiesche E, Epple M. The remineralization of enamel from saliva: a chemical perspective. *Dent J (Basel)*. 2024;12(11):339.

doi: 10.3390/dj12110339

57. Cury JA, Ricomini-Filho AP, Tabchoury CPM. Chapter 5.1: Physicochemical Interactions between Enamel and Oral Fluids. In: Carvalho JC, ed. *Coronal Caries: Evolving Evidence and Clinical Practice*. Vol 31. Monographs in Oral Science. Basel, Switzerland: Karger; 2023:50-61.

doi: 10.1159/000530559

58. Duckworth RM, Jones S. On the relationship between the rate of salivary flow and salivary fluoride clearance. *Caries Res*. 2015;49(2):141-146.

doi: 10.1159/000365949

59. Dziewulska A, Janiszewska-Olszowska J, Bachanek T, Grocholewicz K. Salivary mineral composition in patients with oral cancer. *Magnes Res*. 2013;26(3):120-124.

doi: 10.1684/mrh.2013.0346

60. Kashyap B, Kullaa A. Salivary Metabolites Produced by Oral Microbes in Oral Diseases and Oral Squamous Cell Carcinoma: A Review. *Metabolites*. 2024;14(5):277.

doi: 10.3390/metabo14050277

61. Lacina L, Kolář M, Pfeiferová L, Gál P, Smetana K Jr. Wound healing: insights into autoimmunity, ageing, and cancer ecosystems through inflammation and IL-6 modulation. *Front Immunol*. 2024;15:1403570.

doi: 10.3389/fimmu.2024.1403570

62. Kistner TM, Pedersen BK, Lieberman DE. Interleukin 6 as an energy allocator in muscle tissue. *Nat Metab*. 2022;4(2):170-179.

doi: 10.1038/s42255-022-00538-4

63. Grebenciucova E, VanHaerents S. Interleukin 6: at the interface of human health and disease. *Front Immunol*. 2023;14:1255533.

doi: 10.3389/fimmu.2023.1255533

64. Johnson DE, O'Keefe RA, Grandis JR. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat Rev Clin Oncol*. 2018;15(4):234-248.

doi: 10.1038/nrclinonc.2018.8

65. Rose-John S, Jenkins BJ, Garbers C, Moll JM, Scheller J. Targeting IL-6 trans-signalling: past, present and future prospects. *Nat Rev Immunol*. 2023;23(10):666-681.

doi: 10.1038/s41577-023-00856-y

66. Guo Y, Xie YQ, Gao M, *et al*. Metabolic reprogramming of terminally exhausted CD8+ T cells by IL-10 enhances anti-tumor immunity. *Nat Immunol*. 2021;22(6):746-756.

doi: 10.1038/s41590-021-00940-2

67. Kaplanski G. Interleukin-18: Biological properties and role in disease pathogenesis. *Immunol Rev*. 2018;281(1):138-153.

doi: 10.1111/imr.12616

68. Landy E, Carol H, Ring A, Canna S. Biological and clinical roles of IL-18 in inflammatory diseases. *Nat Rev Rheumatol*. 2024;20(1):33-47.

doi: 10.1038/s41584-023-01053-w

69. Yasuda K, Nakanishi K, Tsutsui H. Interleukin-18 in Health and Disease. *Int J Mol Sci*. 2019;20(3):649.

doi: 10.3390/ijms20030649

70. Somm E, Jornayvaz FR. Interleukin-18 in metabolism: From mice physiology to human diseases. *Front Endocrinol (Lausanne)*. 2022;13:971745.

doi: 10.3389/fendo.2022.971745

71. Bao P, Kodra A, Tomic-Canic M, Golinko MS, Ehrlich HP,

Brem H. The role of vascular endothelial growth factor in wound healing. *J Surg Res*. 2009;153(2):347-358.

doi: 10.1016/j.jss.2008.04.023

72. Perez-Gutierrez L, Ferrara N. Biology and therapeutic targeting of vascular endothelial growth factor A. *Nat Rev Mol Cell Biol*. 2023;24(11):816-834.

doi: 10.1038/s41580-023-00631-w

73. Rovera A, Alzahrani AS, Rovera G, Anderson P. Influence of Parotid Saliva Composition on Phosphate Species' Chemical State in Relation to Dental Caries. *Clin Exp Dent Res*. 2025;11(1):e70089.

doi: 10.1002/cre2.70089

74. Chibly AM, Aure MH, Patel VN, Hoffman MP. Salivary gland function, development, and regeneration. *Physiol Rev*. 2022;102(3):1495-1552.

doi: 10.1152/physrev.00015.2021

75. Mankar N, Kumbhare S, Nikhade P, Mahapatra J, Agrawal P. Role of Fluoride in Dentistry: A Narrative Review. *Cureus*. 2023;15(12):e50884.

doi: 10.7759/cureus.50884

76. Farooq I, Bugshan A. The role of salivary contents and modern technologies in the remineralization of dental enamel: a narrative review. *F1000Res*. 2020;9:171.

doi: 10.12688/f1000research.22499.3

77. Sengupta A, Valdramidou D, Huntley S, Hicks SJ, Carrington SD, Corfield AP. Distribution of MUC1 in the normal human oral cavity is localized to the ducts of minor salivary glands. *Arch Oral Biol*. 2001;46(6):529-538.

doi: 10.1016/s0003-9969(01)00010-3

78. Apostolopoulos V, Stojanovska L, Gargosky SE. MUC1 (CD227): a multi-tasked molecule. *Cell Mol Life Sci*. 2015;72(23):4475-4500.

doi: 10.1007/s00018-015-2014-z

79. Sirviö E, Mikkonen JJW, Koistinen AP, Miinalainen I, Kullaa AM. Localization of transmembrane mucin MUC1 on the apical surface of oral mucosal cells. *Ultrastruct Pathol*. 2019;43(4-5):184-189.

doi: 10.1080/01913123.2019.1687630

80. Kashyap B, Kullaa AM. Regulation of mucin 1 expression and its relationship with oral diseases. *Arch Oral Biol*. 2020;117:104791.

doi: 10.1016/j.archoralbio.2020.104791

81. Grinstead JS, Koganty RR, Krantz MJ, Longenecker BM, Campbell AP. Effect of glycosylation on MUC1 humoral immune recognition: NMR studies of MUC1 glycopeptide-antibody interactions. *Biochemistry*. 2002;41(31):9946-9961.

doi: 10.1021/bi012176z

82. Ployon S, Belloir C, Bonnotte A, Lherminier J, Canon F, Morzel M. The membrane-associated MUC1 improves adhesion of salivary MUC5B on buccal cells. Application to development of an in vitro cellular model of oral epithelium. *Arch Oral Biol*. 2016;61:149-155.

doi: 10.1016/j.archoralbio.2015.11.002

83. Offner GD, Troxler RF. Heterogeneity of high-molecular-weight human salivary mucins. *Adv Dent Res*. 2000;14:69-75.

doi: 10.1177/08959374000140011101

84. Lee LT, Wong YK, Hsiao HY, Wang YW, Chan MY, Chang KW. Evaluation of saliva and plasma cytokine biomarkers in patients with oral squamous cell carcinoma. *Int J Oral Maxillofac Surg*. 2018;47(6):699-707.

doi: 10.1016/j.ijom.2017.09.016

85. Mantovani A, Allavena P, Marchesi F, *et al*. Macrophages as tools and targets in cancer therapy. *Nat Rev Drug Discov*. 2022;21(11):799-820.

doi: 10.1038/s41573-022-00520-5

86. Colvin EK. Tumor-associated macrophages contribute to tumor progression in ovarian cancer. *Front Oncol*. 2014;4:137.

doi: 10.3389/fonc.2014.00137

87. Deshmane SL, Kremlev S, Amini S, Sawaya BE. Monocyte chemoattractant protein-1 (MCP-1): an overview. *J Interferon Cytokine Res*. 2009;29(6):313-326.

doi: 10.1089/jir.2008.0027

88. Shen H, Lee CY, Chen CH. Protein glycosylation as biomarkers in gynecologic cancers. *Diagnostics (Basel)*. 2022;12(12):3177.

doi: 10.3390/diagnostics12123177

89. Bel'skaya LV, Sarf EA, Loginova AI, Vyushkov DM, Choi ED. Potential diagnostic value of salivary tumor markers in breast, lung and ovarian cancer: a preliminary study. *Curr Issues Mol Biol*. 2023;45(6):5084-5098.

doi: 10.3390/cimb45060323

90. Frenkel ES, Ribbeck K. Salivary mucins in host defense and disease prevention. *J Oral Microbiol*. 2015;7:29759.

doi: 10.3402/jom.v7.29759

91. Wang Y, Xu RC, Zhang XL, *et al*. Interleukin-8 secretion by ovarian cancer cells increases anchorage-independent growth, proliferation, angiogenic potential, adhesion and invasion. *Cytokine*. 2012;59(1):145-155.

doi: 10.1016/j.cyto.2012.04.013

92. Chadwick TB, So J, Hertzog PJ, *et al*. Striking the right balance with type I interferon signalling in cancer. *Nat Rev Cancer*. 2026;26:391-414.

doi: 10.1038/s41568-026-00915-1

93. Yu R, Zhu B, Chen D. Type I interferon-mediated tumor immunity and its role in immunotherapy. *Cell Mol Life Sci*. 2022;79(3):191.

doi: 10.1007/s00018-022-04219-z

94. Reddy SU, Sadia FZ, Vancura A, Vancurova I. IFNγ-Induced Bcl3, PD-L1 and IL-8 Signaling in Ovarian Cancer: Mechanisms and Clinical Significance. *Cancers (Basel)*. 2024;16(15):2676.

doi: 10.3390/cancers16152676

95. Ivashkiv LB. IFNγ: signalling, epigenetics and roles in immunity, metabolism, disease and cancer immunotherapy. *Nat Rev Immunol*. 2018;18(9):545-558.

doi: 10.1038/s41577-018-0029-z

96. Browning L, Patel MR, Horvath EB, Tawara K, Jorcyk CL. IL-6 and ovarian cancer: inflammatory cytokines in promotion of metastasis. *Cancer Manag Res*. 2018;10:6685-6693.

doi: 10.2147/CMAR.S179189

97. Wang L, Zang Y, Dejenie R, *et al*. The role of cytokines in ovarian cancer drug resistance. *Int J Cancer*. 2026;158(4):821-834.

doi: 10.1002/ijc.70099

98. Diesch T, Filippi C, Fritschi N, Filippi A, Ritz N. Cytokines in saliva as biomarkers of oral and systemic oncological or infectious diseases: A systematic review. *Cytokine*. 2021;143:155506.

doi: 10.1016/j.cyto.2021.155506

99. Szabo YZ, Slavish DC. Measuring salivary markers of inflammation in health research: A review of methodological considerations and best practices. *Psychoneuroendocrinology*. 2021;124:105069.

doi: 10.1016/j.psyneuen.2020.105069

100. Owen DL, Mahmud SA, Vang KB, *et al*. Identification of Cellular Sources of IL-2 Needed for Regulatory T Cell Development and Homeostasis. *J Immunol*. 2018;200(12):3926-3933.

doi: 10.4049/jimmunol.1800097

101. Moriggl R, Topham DJ, Teglund S, *et al*. Stat5 is required for IL-2-induced cell cycle progression of peripheral T cells. *Immunity*. 1999;10(2):249-259.

doi: 10.1016/s1074-7613(00)80025-4

102. Lin JX, Ge M, Liu CY, *et al*. Tyrosine phosphorylation of both STAT5A and STAT5B is necessary for maximal IL-2 signaling and T cell proliferation. *Nat Commun*. 2024;15(1):7372.

doi: 10.1038/s41467-024-50925-6

103. Chinen T, Kannan AK, Levine AG, *et al*. An essential role for the IL-2 receptor in Treg cell function. *Nat Immunol*. 2016;17(11):1322-1333.

doi: 10.1038/ni.3540

104. Antov A, Yang L, Vig M, *et al*. Essential role for STAT5 signaling in CD25+CD4+ regulatory T cell homeostasis and the maintenance of self-tolerance. *J Immunol*. 2003;171(7):3435-3441.

doi: 10.4049/jimmunol.171.7.3435

105. Seki N, Miyazaki M, Suzuki W, *et al*. IL-4-induced GATA-3 expression is a time-restricted instruction switch for Th2 cell differentiation. *J Immunol*. 2004;172(10):6158-6166.

doi: 10.4049/jimmunol.172.10.6158

106. Linehan LA, Warren WD, Thompson PA, Grusby MJ, Berton MT. STAT6 is required for IL-4-induced germline Ig gene transcription and switch recombination. *J Immunol*. 1998;161(1):302-310.

doi: 10.4049/jimmunol.161.1.302

107. Gascan H, Gauchat JF, Aversa G, Van Vlasselaer P, de Vries JE. Anti-CD40 monoclonal antibodies or CD4+ T cell clones and IL-4 induce IgG4 and IgE switching in purified human B cells via different signaling pathways. *J Immunol*. 1991;147(1):8-13.

doi: 10.4049/jimmunol.147.1.8

108. Daniel B, Czimmerer Z, Halasz L, *et al*. The transcription factor EGR2 is the molecular linchpin connecting STAT6 activation to the late, stable epigenomic program of alternative macrophage polarization. *Genes Dev*. 2020;34(21-22):1474-1492.

doi: 10.1101/gad.343038.120

109. Czimmerer Z, Daniel B, Horvath A, *et al*. The transcription factor STAT6 mediates direct repression of inflammatory enhancers and limits activation of alternatively polarized macrophages. *Immunity*. 2018;48(1):75-90.e6.

doi: 10.1016/j.immuni.2017.12.010

110. Riis JL, Out D, Dorn LD, *et al*. Salivary cytokines in healthy adolescent girls: Intercorrelations, stability, and associations with serum cytokines, age, and pubertal stage. *Dev Psychobiol*. 2014;56(4):797-811.

doi: 10.1002/dev.21149

111. Fiorentino DF, Zlotnik A, Mosmann TR, Howard M, O'Garra A. IL-10 inhibits cytokine production by activated macrophages. *J Immunol*. 1991;147(11):3815-3822.

doi: 10.4049/jimmunol.147.11.3815

112. Ip WKE, Hoshi N, Shouval DS, Snapper S, Medzhitov R. Anti-inflammatory effect of IL-10 mediated by metabolic reprogramming of macrophages. *Science*. 2017;356(6337):513-519.

doi: 10.1126/science.aal3535

113. Sawant DV, Yano H, Chikina M, *et al*. Adaptive plasticity of

IL-10+ and IL-35+ Treg cells cooperatively promotes tumor T cell exhaustion. *Nat Immunol*. 2019;20(6):724-735.

doi: 10.1038/s41590-019-0346-9

114. Saraiva M, O'Garra A. The regulation of IL-10 production by immune cells. *Nat Rev Immunol*. 2010;10(3):170-181.

doi: 10.1038/nri2711

115. Agarwal A, Tressel SL, Kaimal R, *et al*. Identification of a metalloprotease-chemokine signaling system in the ovarian cancer microenvironment: implications for antiangiogenic therapy. *Cancer Res*. 2010;70(14):5880-5890.

doi: 10.1158/0008-5472.CAN-09-4341

116. Kulbe H, Chakravarty P, Leinster DA, *et al*. A dynamic inflammatory cytokine network in the human ovarian cancer microenvironment. *Cancer Res*. 2012;72(1):66-75.

doi: 10.1158/0008-5472.CAN-11-2178

117. Coward J, Kulbe H, Chakravarty P, *et al*. Interleukin-6 as a therapeutic target in human ovarian cancer. *Clin Cancer Res*. 2011;17(18):6083-6096.

doi: 10.1158/1078-0432.CCR-11-0945

118. Deng W, Gu X, Lu Y, *et al*. Down-modulation of TNFSF15 in ovarian cancer by VEGF and MCP-1 is a pre-requisite for tumor neovascularization. *Angiogenesis*. 2012;15(1):71-85.

doi: 10.1007/s10456-011-9244-y

119. Chauhan SC, Kumar D, Jaggi M. Mucins in ovarian cancer diagnosis and therapy. *J Ovarian Res*. 2009;2:21.

doi: 10.1186/1757-2215-2-21

120. Singh AP, Senapati S, Ponnusamy MP, *et al*. Clinical potential of mucins in diagnosis, prognosis, and therapy of ovarian cancer. *Lancet Oncol*. 2008;9(11):1076-1085.

doi: 10.1016/S1470-2045(08)70277-8

121. Stähli C, Caravatti M, Aeschbacher M, Kocyba C, Takacs B, Carmann H. Mucin-like carcinoma-associated antigen defined by three monoclonal antibodies against different epitopes. *Cancer Res*. 1988;48(23):6799-6802.

122. Bhogadia M, Edgar M, Hunwin K, Page G, Grootveld M. Detection and Quantification of Ammonia as the Ammonium Cation in Human Saliva by 1H NMR: A Promising Probe for Health Status Monitoring, with Special Reference to Cancer. *Metabolites*. 2023;13(7):792.

doi: 10.3390/metabo13070792

123. Fasoulakis Z, Koutras A, Ntounis T, *et al*. Ovarian Cancer and Glutamine Metabolism. *Int J Mol Sci*. 2023;24(5):5041.

doi: 10.3390/ijms24055041

124. Boldyrev AA, Aldini G, Derave W. Physiology and pathophysiology of carnosine. *Physiol Rev*. 2013;93(4):1803-1845.

doi: 10.1152/physrev.00039.2012

125. Hartman PE, Hartman Z, Ault KT. Scavenging of singlet molecular oxygen by imidazole compounds: high and sustained activities of carboxy terminal histidine dipeptides and exceptional activity of imidazole-4-acetic acid. *Photochem Photobiol*. 1990;51(1):59-66.

doi: 10.1111/j.1751-1097.1990.tb01684.x

126. Wong LW, Mak SH, Goh BH, Lee WL. The Convergence of FTIR and EVs: Emergence Strategy for Non-Invasive Cancer Markers Discovery. *Diagnostics (Basel)*. 2022;13(1):22.

doi: 10.3390/diagnostics13010022

127. Yap T, Pruthi N, Seers C, Belobrov S, McCullough M, Celentano A. Extracellular Vesicles in Oral Squamous Cell Carcinoma and Oral Potentially Malignant Disorders: A Systematic Review. *Int J Mol Sci*. 2020;21(4):1197.

doi: 10.3390/ijms21041197

128. McNab AI, Heinze T, McCue AJ, Dionisi D, Anderson JA. Quantification of hydrocarbon species on surfaces by combined microbalance-FTIR. *Spectrochim Acta A Mol Biomol Spectrosc*. 2017;181:65-72.

doi: 10.1016/j.saa.2017.03.030

129. Bozkurt-Girit O, Kilic MA. FTIR spectroscopic characterization reveals short-term macromolecular responses to photobiomodulation in mesenchymal stem cells. *Sci Rep*. 2025;15:31051.

doi: 10.1038/s41598-025-15190-7

130. Bano A, Vats R, Verma D, Yadav P, Kamboj M, Bhardwaj R. Exploring salivary exosomes as early predictors of oral cancer in susceptible tobacco consumers: noninvasive diagnostic and prognostic applications. *J Cancer Res Clin Oncol*. 2023;149(17):15781-15793.

doi: 10.1007/s00432-023-05343-4

131. Zlotogorski-Hurvitz A, Dekel BZ, Malonek D, Yahalom R, Vered M. FTIR-based spectrum of salivary exosomes coupled with computational-aided discriminating analysis in the diagnosis of oral cancer. *J Cancer Res Clin Oncol*. 2019;145(3):685-694.

doi: 10.1007/s00432-018-02827-6

132. Schneider C. An update on products and mechanisms of lipid peroxidation. *Mol Nutr Food Res*. 2009;53(3):315-321.

doi: 10.1002/mnfr.200800131

133. Kim RS, LaBella FS. Comparison of analytical methods for monitoring autoxidation profiles of authentic lipids. *J Lipid Res*. 1987;28(9):1110-1117.

doi: 10.1016/S0022-2275(20)38624-7

134. Adeoye J, Su YX. Artificial intelligence in salivary biomarker discovery and validation for oral diseases. *Oral Dis*. 2024;30:23-37.

doi: 10.1111/odi.14641

135. Koundouros N, Poulogiannis G. Reprogramming of fatty acid metabolism in cancer. *Br J Cancer*. 2020;122:4-22.

doi: 10.1038/s41416-019-0650-z

136. Cristaldi M, Mauceri R, Di Fede O, *et al*. Salivary biomarkers for oral squamous cell carcinoma diagnosis and follow-up: current status and perspectives. *Front Physiol*. 2019;10:1476.

doi: 10.3389/fphys.2019.01476

137. Zhang Y, Liu J, Liu S, *et al*. Extracellular vesicles in oral squamous cell carcinoma: current progress and future prospect. *Front Bioeng Biotechnol*. 2023;11:1149662.

doi: 10.3389/fbioe.2023.1149662

138. Ayala A, Muñoz MF, Argüelles S. Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. *Oxid Med Cell Longev*. 2014;2014:360438.

doi: 10.1155/2014/360438

139. Dixson AC, Dawson TR, Di Vizio D, *et al*. Context-specific regulation of extracellular vesicle biogenesis and cargo selection. *Nat Rev Mol Cell Biol*. 2023;24(7):454-476.

doi: 10.1038/s41580-023-00576-0

140. Agarwal A, Rani M, Saha GK, *et al*. Disregulated expression of the Th2 cytokine gene in patients with intraoral squamous cell carcinoma. *Immunol Invest*. 2003;32(1-2):17-30.

doi: 10.1081/imm-120019205

141. Sutera S, Furchi OA, Pentenero M. Investigating tumor-infiltrating lymphocytes in the microenvironment of oral squamous cell carcinoma and oral potentially malignant disorders: can they shift our perspective? A scoping review. *J Clin Med*. 2025;14(2):606.

doi: 10.3390/jcm14020606

142. Kondoh N, Mizuno-Kamiya M. The role of immune modulatory cytokines in the tumor microenvironments of head and neck squamous cell carcinomas. *Cancers (Basel)*. 2022;14(12):2884.

doi: 10.3390/cancers14122884

143. Jin J, Lin J, Xu A, *et al*. CCL2: An Important Mediator Between Tumor Cells and Host Cells in Tumor Microenvironment. *Front Oncol*. 2021;11:722916.

doi: 10.3389/fonc.2021.722916

144. Yoshimura T. The chemokine MCP-1 (CCL2) in the host interaction with cancer: a foe or ally? *Cell Mol Immunol*. 2018;15(4):335-345.

doi: 10.1038/cmi.2017.135

145. Ferrari E, Pezzi ME, Cassi D, Pertinhez TA, Spisni A, Meleti M. Salivary Cytokines as Biomarkers for Oral Squamous Cell Carcinoma: A Systematic Review. *Int J Mol Sci*. 2021;22(13):6795.

doi: 10.3390/ijms22136795

146. Pan K, Li Q, Guo Z, Li Z. Healing action of interleukin-4 in acute and chronic inflammatory conditions: mechanisms and therapeutic strategies. *Pharmacol Ther*. 2025;265:108760.

doi: 10.1016/j.pharmthera.2024.108760

147. Van Dyken SJ, Locksley RM. Interleukin-4- and interleukin-13-mediated alternatively activated macrophages: roles in homeostasis and disease. *Annu Rev Immunol*. 2013;31:317-343.

doi: 10.1146/annurev-immunol-032712-095906

148. Boutilier AJ, Elsawa SF. Macrophage Polarization States in the Tumor Microenvironment. *Int J Mol Sci*. 2021;22(13):6995.

doi: 10.3390/ijms22136995

149. Zhang Q, Sioud M. Tumor-Associated Macrophage Subsets: Shaping Polarization and Targeting. *Int J Mol Sci*. 2023;24(8):7493.

doi: 10.3390/ijms24087493

150. Ambudkar IS. Polarization of calcium signaling and fluid secretion in salivary gland cells. *Curr Med Chem*. 2012;19(34):5774-5781.

doi: 10.2174/092986712804143321

151. Ambudkar IS. Ca^{2+} signaling and regulation of fluid secretion in salivary gland acinar cells. *Cell Calcium*. 2014;55(6):297-305.

doi: 10.1016/j.ceca.2014.02.009

152. Ambudkar I. Calcium signaling defects underlying salivary gland dysfunction. *Biochim Biophys Acta Mol Cell Res*. 2018;1865(11):1771-1777.

doi: 10.1016/j.bbamcr.2018.07.002