BMC Cancer

RESEARCH Open Access

Identification of novel serum lipid metabolism potential markers and metabolic pathways for oral cancer: a population-based study

Na Wang[1,2†], Yujia Chen[1,2†], Jianli Lin[1,2†], Yulan Lin[1,2], Haoyuan Song[1,2], Weihai Huang[1,2], Liling Shen[1,2], Fa Chen[1,2], Fengqiong Liu[1,2], Jing Wang[3], Yu Qiu[4], Bin Shi[4], Ling Li[5], Lisong Lin[4], Lizhen Pan[6*] and Baochang He[1,2*]

Abstract

Objective This study aims to identify potential lipid biomarkers and metabolic pathways associated with oral cancer (OC). Then to establish and evaluate disease classification models capable of distinguishing OC patients from healthy controls.

Methods A total of 41 OC patients and 41 controls were recruited from a hospital in Southeast China to examine the serum lipidomics by Ultra-high Performance Liquid Chromatography Q Exactive Mass Spectrometry (UHPLC-QE-MS).

Results The total serum lipid profile showed that triglycerides accounted for the highest proportion of total metabolites, reaching 35.90% of the total. A total of 74 different metabolites were screened (12 up-regulated and 62 down-regulated), mainly enriched in the glycerophospholipid metabolism pathway. The three most significant changes in lipid metabolites were phosphatidylcholine (PC(18:3e/17:2)), acylcarnitine (ACar(14:2)), and glucuronosyldiacylglycerol (GlcADG(14:1/14:1)). The disease classification model, constructed using a KNN algorithm with 13 metabolites selected through LASSO screening, achieved the best performance, with an AUC of 0.978 (0.955-1.000).

Conclusion Lipid metabolic biomarkers identified in this study exhibit potential as candidate biomarkers for OC diagnosis. Further validation through prospective studies is required to confirm their clinical utility in early detection.

Keywords Oral cancer, Lipid metabolism, Case-control study, Serum, Biomarkers

†Na Wang, Yujia Chen and Jianli Lin these authors have contributed equally to this work.

*Correspondence:
Lizhen Pan
lilylilylily1234@163.com
Baochang He
prof_hbc@163.com
[1]Department of Epidemiology and Health Statistics, Fujian Provincial Key Laboratory of Environment Factors and Cancer, School of Public Health, Fujian Medical University, 1 Xueyuan Road, Fujian 350108, China
[2]Key Laboratory of Ministry of Education for Gastrointestinal Cancer, Fujian Medical University, Fujian, China
[3]Laboratory Center, School of Public Health, The Major Subject of Environment and Health of Fujian Key Universities, Fujian Medical University, Fujian, China
[4]Department of Oral and Maxillofacial Surgery, the First Affiliated Hospital of Fujian Medical University, Fuzhou, China
[5]International Nursing School, Hainan Medical University, Hainan, China
[6]Department of Nursing, the First Affiliated Hospital of Fujian Medical University, 88 Jiaotong Road, Fujian 350004, China

Introduction

Oral cancer (OC) is a malignant tumor occurred in lips, tongue, gingiva, floor of mouth, palate, and buccal mucosa [1]. While the pathological biopsy is the main definitive method of diagnosing OC with high specificity, the high invasive limited its utility in early screening [2]. Therefore, more attention is needed to create a more accurate, sensitive, cost-effective, and non-invasive test. Molecular markers from body fluid including saliva, urine, or others for detecting cancer and monitoring disease progression have been emphasized recently [3]. Blood or serum-based tests offer the aforementioned merits and are repeatable [4].

Blood lipids are total cholesterol (TC), triglycerides (TAG), and phospholipids et al. in serum. TC, which serves as the predominant constituent of cell and organelle membranes and the source of energy, playing an necessary role in biological processes (including cell proliferation and tissue division) [5]. The early accumulation of TAGs is associated with Fas-induced apoptosis [6]. Phospholipids play important roles in transformation, cancer progression, and metastasis by undergoing cellular signaling, chemical-energy storage, and cell-cell interactions in tissues [7]. Previous studies revealed that lipid metabolism is closely associated with cancers and the activation of oncogenic pathways accelerates lipid synthesis [8]. Dysregulation in lipid metabolism is also one of the most significant metabolic alterations in cancer [9].

Untargeted metabolomics is an unbiased analysis that can systematically examine metabolic characteristics in various contexts. Lipidomics, a major branch of metabolomics, is extensively used to test the total lipids in biological systems. This high-throughput technology explores the changes of lipid metabolic profiles and can select biomarkers with high accuracy, sensitivity, and specificity to improve diagnostic accuracy [10]. With the advent of high-throughput data, the precision of feature selection and model construction has become crucial. LASSO (Least Absolute Shrinkage and Selection Operator) and RFE (Recursive Feature Elimination) are two commonly employed methods for variable selection, widely applied in cancer-related prediction models. LASSO introduces an L1 regularization term that penalizes regression coefficients, effectively selecting key variables while reducing overfitting [11]. RFE operates by recursively eliminating less important variables to improve model performance [12]. This method evaluates feature importance and iteratively refines the model to retain only the most significant predictors [13]. After variable selection, algorithms such as K-Nearest Neighbors (KNN) and recursive partitioning are frequently used to construct disease prediction models. Many diseases diagnostic models established by KNN algorithms have achieved AUC values exceeding 0.8, demonstrating robust performance in differentiating disease cases from controls [14–16]. And recursive partitioning models always with highly accuracy and reliability in disease prediction [17, 18]. The integration of these approaches has garnered significant attention in the development of cancer-related predictive models due to their ability to enhance predictive accuracy.

In our study, an Ultra-high Performance Liquid Chromatography Q Exactive Mass Spectrometry (UHPLC-QE-MS) based technique was performed to untargeted examine the total lipids present in serum to identify the potential lipid metabolic biomarkers of OC. And then to establish and evaluate disease classification models capable of distinguishing OC patients from healthy controls using different feature selection and model construction methods.

Materials and methods

Study design and study population

This case-control study was carried out in the First Affiliated Hospital of Fujian Medical University (Fuzhou city of Fujian province, China) from December 2020 to August 2022. Patients were recruited if: (a) histologically confirmed primary OC; (b) has lived in Fujian Province for at least ten years. We excluded patients if they: (a) with recurrent OC, systemic inflammatory diseases (such as hepatitis, nephritis, and long-term immunosuppressive, et al.); (b) received prior care that might affect serum lipid metabolism (such as chemotherapy, radiation, and lipid-regulating drugs, et al.). At the same time of cases recruited, healthy controls were randomly admitted from the physical examination center in the same hospital. Controls were excluded if: (a) has lived in Fujian Province for less than ten years; (b) has the history (or family history) of cancers; (c) with systemic inflammatory diseases (such as hepatitis, nephritis, and long-term immunosuppressive, et al.); (d) received prior care that might affect serum lipid metabolism. Finally, 41 patients and 41 healthy controls were recruited.

Data collection and definition

Basic information of patients including age, gender, weight, height, place of residence, tobacco, alcohol, and tea consumption, hypertension, and diabetes was collected via an interview-based structured questionnaire. This questionnaire was produced by the University of Utah and the Institute of Oncology, the Chinese Academy of Medical Sciences, in collaboration with renowned Chinese epidemiologists, and was tailored to reflect the sociocultural and environmental circumstances specific to China [19]. The questionnaire can be found in supplement material.

Age values were grouped into two groups (< 60 and ≥ 60 years old). BMI values were categorized according to the Chinese definitions (< 18.5, 18.5–23.9, ≥ 24.0 kg/m^2).

Serum preparation

Venous blood (5 ml) of all participants was collected the day after diagnosis or physical examination. The serum was rapidly separated and stored at -80℃ immediately. 100 µl of serum after thawing from each sample was taken and mixed with 480 µl of extraction solvent ($V_{MTBE}/V_{methanol}$=5:1, including Isopropanol). After vortexing for 30s and low-temperature ultrasonication for 10 min, the mixture was stored at a temperature of -40℃ for 1 h. Then after centrifuging the mixture at 3000 rpm for 15 min to remove the denatured protein, 300 µl of the upper layer of the mixture was transferred and dried. Dried samples were redissolved in 100 µl of dissolving solution ($V_{dichloromethane}/V_{methanol}$ = 1:1), after vortexing for 30s, low-temperature ultrasonication for 10 min, and centrifuging the mixture at 13,000 rpm for 15 min, 75 µl samples were taken and analyzed by UHPLC-QE-MS.

Furthermore, 20 µl of each sample were mixed as a quality control (QC) sample to examine the stability of instrumental analysis. The handling and analysis of QC samples matched those of real samples exactly.

Metabolomics analysis

The UHPLC (Dionex) coupled to the Q-Exactive Orbitrap MS (Thermo Fisher Scientific) was performed for metabolomic analysis of serum samples. A Phenomen Kinetex C18 (2.1*100 mm, 1.7 µm) column was performed for lipids separation. The mobile phase A was acetonitrile/water (60:40, v/v) with 10 mM ammonium formate contained in, while mobile phase B was isopropanol/acetonitrile (90:10, v/v) with 10 mM ammonium formate contained in. The detailed settings of gradient parameters and instrument conditions were shown in Supplement Tables 1 and 2.

The raw data was collected by the Xcalibur software, transformed to the mzXML format by Proteo Wizard MS converter, and processed for peak alignment, retention time correction, and peak area extraction by XCMS software. Metabolite annotation was performed using an in-house MS2 database (BiotreeDB) with a cutoff score of 0.3.

Statistical analysis

Baseline characteristics were assessed for comparability using the Chi-square test, while metabolite comparisons (after normalized) utilized the student's t-test. Principal Component Analysis (PCA) was conducted to assess data distribution and sample classification based on log-transformed and unit variance (UV) scaled metabolite concentrations. Then, the Orthogonal Projections to Latent Structures-Discriminant Analysis (OPLS-DA) was used to examine the changes of each metabolic between case and control groups based on the log-transformed plus UV scaled data, the 7-fold cross-validation was applied to assess model validity and a permutation test with 200 replacements was performed to check for overfitting. Metabolites with Variable Importance in Projection (VIP) > 1 in OPLS-DA and FDR adjusted $P < 0.05$ in student's t-test were selected for functional enrichment analysis of disease-related dominant pathways using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Key metabolites influencing oral cancer were selected by the least absolute shrinkage and selection operator (LASSO) or Recursive Feature Elimination (RFE), and four disease-classification models were then established by combining variable selection methods with K-Nearest Neighbor (KNN) or recursive partitioning algorithms. The Receiver operating characteristic (ROC) curves and the area under the curves (AUC) were used to evaluate the prediction efficacy. The relationship between metabolites in the most efficient model and oral cancer was further examined using logistic regression.

The PCA and OPLS-DA analyses were performed by SIMCA14.1 (Umetrics. Sweden), while the other statistical analyses were performed by R software (version 4.1.3). Two-tailed P value less than 0.05 was regarded as statistically significant.

Results

Patient characteristics

The baseline characteristics of 41 OC patients and 41 healthy controls in this study were provided in Table 1. A total of 43 males and 39 females were included with a mean age of 57.45 ± 12.99 years old. Besides hypertension and diabetes, the distributions of participants in other characteristics including gender, age, BMI, smoking, alcohol, and tea drinking were similar ($P > 0.05$).

Evaluation of the stability and reproducibility of the instrument analysis

The Supplement Fig. 1 showed the quality control results, the retention time and peak area overlapped well in total ion chromatography (TIC) of QC samples, indicating great instrument stability. The extracted ion chromatogram (EIC) showed the retention time and peak area of internal standard in QC samples overlapped well, indicating the great stability of data collection for the instrument. The TIC diagram of blank sample was showed in Supplement Fig. 2, no discernible chromatographic peak was observed in the blank sample for most of the time, indicating the effective control of the residue and the level of contamination among samples was acceptable.

Lipid profiling of lipid species

Serum lipid profiles were identified by lipidomics from a total of 92 serum samples (41 oral cancer patients, 41 healthy controls, and 10 QC samples). Using UHPLC-QE-MS methods, a total of 20,258 peaks were extracted

Table 1 General demographic characteristics of the case and control groups

Variables	Case[*n*(%)] (*n* = 41)	Control[*n*(%)] (*n* = 41)	χ^2	*P*
Age (year)			0.05	0.825
< 60	20(48.78)	19(46.34)		
≥ 60	21(51.22)	22(53.66)		
Gender			0.05	0.825
Male	22(53.66)	21(51.22)		
Female	19(46.34)	20(48.78)		
BMI (kg/m^2)			1.85	0.397
18.5–23.9	23(56.10)	18(43.90)		
< 18.5	5(12.20)	4(9.76)		
≥ 24	13(31.71)	19(46.34)		
Hypertension			4.20	0.040
No	30(73.17)	21(51.22)		
Yes	11(26.83)	20(48.78)		
Diabetes			4.48	0.034
No	31(75.61)	38(92.68)		
Yes	10(24.39)	3(7.32)		
Smoke			2.82	0.093
No	25(60.98)	32(78.05)		
Yes	16(39.02)	9(21.95)		
Alcohol drink			0.82	0.364
No	33(80.49)	36(87.80)		
Yes	8(19.51)	5(12.20)		
Tea drink			< 0.01	> 0.999
No	27(65.85)	27(65.85)		
Yes	14(34.15)	14(34.15)		

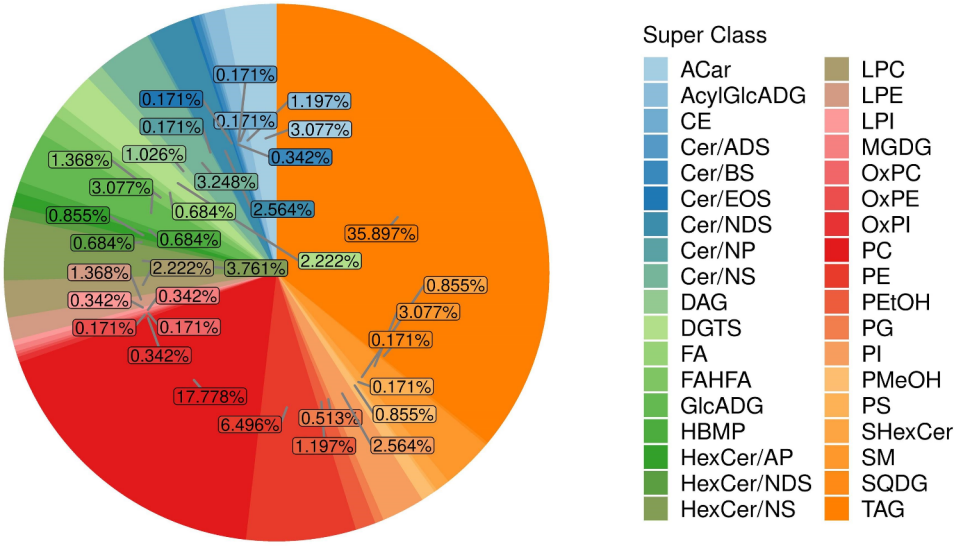

Fig. 1 Metabolite classification and proportion pie chart. The different color blocks in the figure represent different classification categories, and the percentage represents the percentage of metabolites belonging to this type in all identified metabolites

and 12,548 peaks of them were finally remained after pretreatment. Overall, a total of 585 metabolites in 29 lipid classes (including sphingomyelin (SM), triglyceride (TG), phosphatidylethanolamine (PE), and phosphatidylinositol (PI), et al.) were identified. A heat map representing the hierarchical clustering analysis of lipid species across all samples was shown in Supplement Fig. 3.

The identified metabolites were categorized based on their classification. According to the proportion pie chart figure (Fig. 1), TAG accounted for the highest proportion

of metabolites (35.90%), followed by Phosphatidylcholine (PC) and Phosphatidylethanolamine (PE) (17.78% and 6.50%, respectively). These three metabolites together account for above 60% of the total metabolites.

Principal component and orthogonal partial least squares discriminant analysis of cases and controls

The first two principal components (PC), with loading value greater than 1, namely PC1 and PC2, were derived from the PCA analysis (Supplementary Table 1). A PCA model constructed using these two PCs demonstrated that only some samples showed a clear trend of separation between the case and control groups (Supplementary Fig. 4).OPLS-DA was performed to discriminate the metabolites that were responsible for the separation. As shown in Fig. 2A, an obvious separation existed between the cases and controls. For the OPLS-DA model, the values of R2X, R2Y, and Q2 were 0.217, 0.765, and 0.561, respectively, indicating good performance of the model. The results of 200 permutation tests demonstrated that the OPLS-DA model did not over-fit (Fig. 2B).

Screening of differential metabolites and functional enrichment analysis

A total of 74 differential metabolites were identified through the analysis described above. As showed in Supplement Fig. 5, compared with the control group, 12 metabolites (including PC, PE, Glucuronosyldiacylglycerol (GlcADG), and Diacylglyceryl trimethylhomoserine (DGTS), et al.) were up-regulated, and 62 metabolites (including TAG, Sphingomyelin (SM), Acylcarnitine (ACar), and Free Fatty Acid (FA), et al.) were down-regulated. Table 2 showed the top 10 most significantly up-regulated or down-regulated metabolites. The box plot visualized the difference of the 20 metabolites between case and control groups (Fig. 3), and the match rod diagram showed the 2 logarithm of fold-change (Log_2FC) for each metabolite (Supplement Fig. 6). The KEGG functional enrichment analysis indicates that 74 differential metabolites enriched in 14 differential metabolite pathways (Supplement Table 2). The most relative pathway is Glycerophospholipid metabolism, which significantly down-regulated in patients with oral cancer. The following are Linoleic acid metabolism and alpha-Linolenic acid metabolism, and the least are Glycerolipid metabolism and Arachidonic acid metabolism (Supplement Fig. 7).

Construction of the diagnosed model based on the metabolites

We use LASSO regression and RFE analysis to select the differential metabolites and construct disease classification model, RFE use three methods for input variable selection including random forest (RF), BaggingClassifie, and Support Vector Machine (SVM). Based on the λ with one standard error (1se) value, 13 metabolites are finally selected by LASSO regression (Supplementary Fig. 8). The comparison of RFE strategy accuracy for three different variable selection methods is shown in Supplement Fig. 9, RF based RFE strategy including 14 metabolites with the highest accuracy and consistency is finally selected.

The KNN and recursive partitioning algorithm are used to construct disease classification models based on 13 metabolites (selected by LASSO) and 14 metabolites (selected by RF-RFE). As shown in Supplement Table 3, the model constructed by KNN algorithm with metabolites selected by LASSO has a great predictive performance (AUC: 0.978, 95%CI: 0.955-1.000) (Supplement Fig. 10). Moreover, univariate and multivariate Logistic regression are used to explore the association between those metabolites and oral cancer. As shown in Table 3, elevated levels of metabolites such as PE(18:5e/21:2), PC(22:4e/22:6), and GlcADG(14:1/14:1) may be associated with an increased risk of oral cancer, while reduced levels of metabolites including PC(18:3e/17:2), PC(17:0/20:3), TAG(16:0/20:4/20:4), ACar(18:2), ACar(16:0),TAG(18:1/20:4/22:4), PC(16:2/26:4), PEtOH(24:4/22:6), PC(20:4/20:5), and TAG(13:0/13:0/21:2) may be associated with a decrease risk (all $P<0.05$). The results are still obvious after adjusting hypertension and diabetes (all $P<0.05$).

Discussion

Here, we conducted a case-control study involving 41 OC patients and 41 healthy controls from a hospital in southern China. After checking the serum levels of lipid metabolites for participants by UHPLC-QE-MS, our results revealed that PC(18:3e/17:2) and ACar(14:2) were among the most significantly down-regulated metabolites, while GlcADG(14:1/14:1) was the most significantly up-regulated metabolite. Notably, the primary metabolic alterations between OC patients and healthy controls were associated with glycerophospholipid metabolism. The disease classification model constructed by KNN algorithm, incorporating 13 metabolites selected via LASSO regression, demonstrated strong performance in accurately identifying oral cancer patients.

Lipids play important roles in many cellular processes including membrane construction, energy storage and supply, transduction of cellular signals, redox homeostasis, and protein modification, all of which are considered the core of the occurrence and development of cancer [20, 21]. Reprogramming of lipid metabolism has been newly recognized as a hallmark of malignancy [22]. Our study found that some metabolites (including PC(22:4e/22:6), PC(2:0/26:2), PC(10:0/22:3), and PC(14:1e/18:5)) were found to be up-regulated in serum of oral cancer

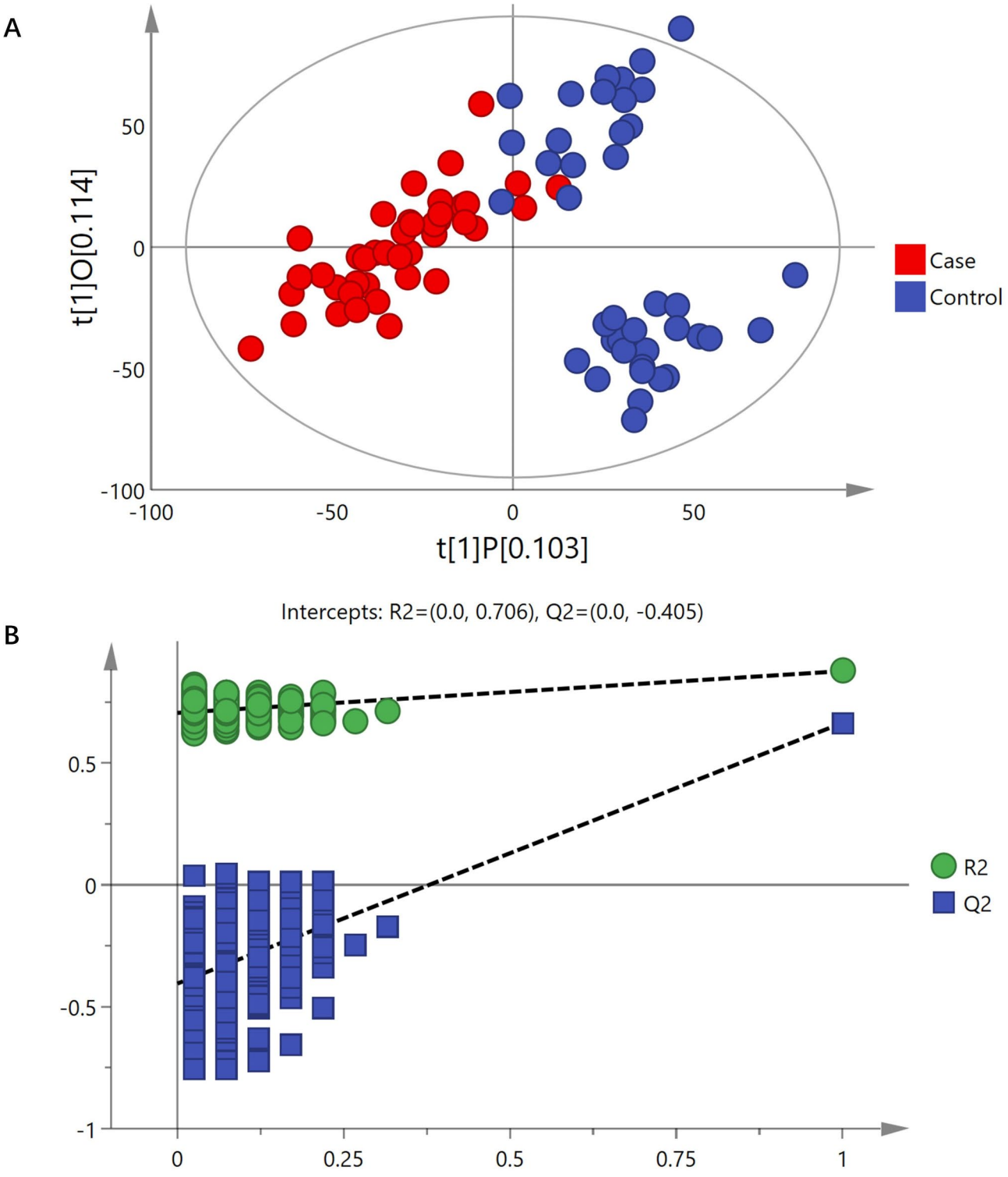

Fig. 2 Scatter chart of OPLS-DA model scores between the case group and the controls (**A**) and Results of 200 replacement tests for OPLS-DA models in the case group and the controls (**B**). Note: **A**: The abscissa t [1]P represents the predicted principal component score of the first principal component, showing the differences between the two groups; The ordinate t [1]O represents the orthogonal principal component score, showing the differences between samples within the group. **B**: The abscissa represents the substitution retention degree of the substitution test, the ordinate represents the value of R^2 or Q^2, the green dot represents the R^2 value obtained from the substitution test, the blue square dot represents the Q^2 value obtained from the substitution test, and the two dashed lines represent the regression lines of R^2 and Q^2, respectively

Table 2 Screening results of differential mMetabolites (partial)

MS2 name	rtime	mz	type	MEAN (case)	MEAN (control)	VIP	*P*-value	Q-value	Log_2 Fold Change
UP									
PC(22:4e/22:6)	351.522	868.602	POS	2.264	1.086	1.761	< 0.001	0.001	1.060
GlcADG(14:1/14:1)	321.437	702.482	POS	6.775	3.170	1.948	0.001	0.003	1.096
PC(2:0/26:2)	354.599	674.472	POS	3.208	1.731	1.697	0.001	0.003	0.890
PE(18:5e/21:2)	378.684	762.529	POS	1.536	0.811	1.838	0.001	0.004	0.921
PC(10:0/22:3)	378.716	728.521	POS	0.595	0.326	1.731	0.002	0.007	0.867
DGTS(2:0/26:1)	189.447	654.517	POS	0.252	0.151	1.413	0.003	0.007	0.736
DGTS(3:0/21:1)	66.384	598.454	POS	9.039	5.333	1.645	0.003	0.008	0.761
PC(14:1e/18:5)	92.241	708.491	POS	0.351	0.202	1.499	0.005	0.012	0.799
AcylGlcADG(12:0/18:3/16:4)	363.912	958.665	POS	5.897	4.035	1.455	0.008	0.015	0.547
TAG(12:2/12:2/15:3)	378.177	689.487	POS	3.391	2.326	1.390	0.008	0.016	0.544
DOWN									
SM(d14:1/20:2)	336.138	699.546	POS	0.318	0.553	1.183	< 0.001	0.001	-0.799
ACar(16:0)	120.649	400.343	POS	3.836	5.812	1.970	< 0.001	0.001	-0.599
TAG(12:3/12:3/14:3)	163.289	671.442	POS	0.257	0.423	1.882	< 0.001	0.001	-0.721
ACar(16:2)	78.605	396.312	POS	0.222	0.410	1.941	< 0.001	0.002	-0.887
FA(20:4)	115.349	303.233	NEG	4.824	7.656	2.000	< 0.001	0.002	-0.666
TAG(12:3/18:5/18:5)	407.982	803.540	POS	4.259	6.228	1.500	< 0.001	0.003	-0.548
ACar(16:1)	92.239	398.328	POS	0.742	1.196	2.155	< 0.001	0.003	-0.688
ACar(12:1)	63.078	342.265	POS	0.438	0.797	1.823	0.001	0.003	-0.865
ACar(14:2)	65.798	368.281	POS	0.536	1.057	1.872	0.001	0.003	-0.981
ACar(18:2)	97.247	424.344	POS	3.061	4.199	1.470	0.001	0.003	-0.456

Note: MS2 name is the name of the substance obtained through qualitative matching analysis by secondary mass spectrometry; rtime, the chromatographic retention time of the substance; mz, mass charge ratio of the characteristic ions of the substance; type, collection mode (POS is positive ion, NEG is negative ion); MEAN, the mean value of the relative quantitative value of the substance (the case group vs. the control group); VIP, the projected importance of the variable obtained by the substance in the OPLS-DA model; *P*-value, the P-value obtained by the student's t test for the substance; *Q*-value, *P* value after multiple correction of FDR; Log_2 Fold Change, the multiple relationship between the two groups of the substance (logarithm based on 2)

patients. This finding is consistent with a previous study that observed serum levels of phosphatidylcholines (PCs) increasing in esophageal adenocarcinoma [23]. Besides that, the abundances of several PCs (including 32:0, 32:1, 34:1 and 36:3) are significantly higher in papillary thyroid cancerous tissue compared with non-cancerous tissue [24]. PCs are the major phospholipid in most eukaryotic cells which plays fundamental structural and regulatory function in cellular membranes [25]. The increasing level of PCs may indicate tumor invasion. The PCs are generated by the phosphorylation of choline which catalyzed by choline kinase (Chk), while the increasing level of the Chk-a gene in cancers has been confirmed [26, 27].

Acylcarnitines (ACar) are fatty acid metabolites with the function of transferring the acyl groups from the cytosol into the mitochondrial matrix. The acyl groups then will undergo β-oxidation, producing energy to maintain cell activity [28]. Thus, the consumption of ACar is closely associated with the activity of β-oxidation. Interestingly, a previous study has already found that fatty acid β-oxidation is a highly represented metabolic process in a proteomics analysis of human renal cell carcinoma tissue [29]. The increased fatty acid β-oxidation was also observed in many cancers including prostate cancer and colon cancer [30, 31]. The serum levels of some ACars (such as ACar (12:1), ACar (16:1), and ACar (14:2), et al.) were significantly down-regulated in oral cancer patients in our study, such negative association was found in patients with lung cancer [32]. Furthermore, after testing the serum level of ACar in 54 cancer patients and 81 non-cancer controls, Dileep S et al. found that patients have significantly lower serum levels of ACar compared with controls (6.7 vs. 11.5 nmol/ml) [33].

Glycolipids containing α-linked glucuronic or galacturonic acid can activate NKT cells, playing a key role in body immunity [34]. Previous studies also supported that the anionic glycolipids related to glucuronosyldiacylglycerol (GlcADG) inhibit protein kinase Akt, and Akt is a key regulator of cancer cells proliferation and survival [35]. We found the serum level of GlcADG significantly down-regulated in oral cancer patients, and this result is rarely reported in other cancers.

The KEGG analysis results showed that differential lipid metabolites mainly enriched in glycerophospholipid metabolism pathway, indicating that the perturbation of glycerophospholipid metabolism was highly associated to OC. Yang et al. also identified that glycerophospholipid metabolism was related with esophageal squamous cell cancer tumorigenesis and progression [36]. By testing the lipid metabolites profile in tissue for 40 lung squamous cell carcinoma patients, Zeng et al. found that glycerophospholipid metabolism changes most in tumor tissues

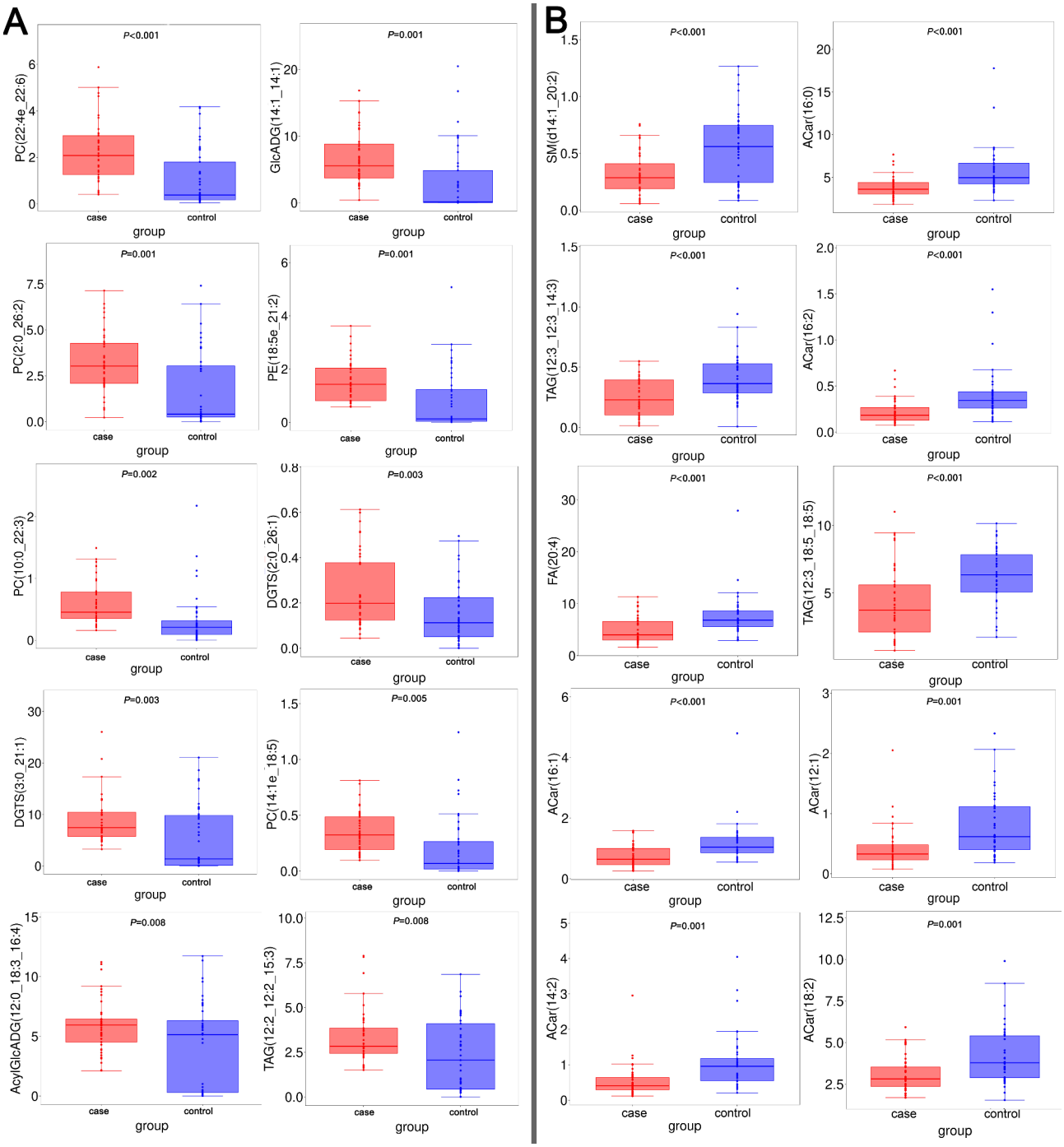

Fig. 3 Box chart for relative quantitative comparison of differential metabolites between the case group and the control group. Note: The P-value in the figure shows the results of student's t test. The first two columns from the left (**A**) show metabolites that are significantly upregulated in the case group. The last two columns (**B**) are significantly downregulated metabolites in the case group

among eight differential lipid metabolism pathways [37]. Furthermore, a previous tissue lipidomics study by Dickinson A et al. found that the glycerophospholipid metabolism pathway containing the greatest number of changed lipid species in OSCC patients [38]. But this study only contained 5 patients with primary oral cancer, indicating the result should undergo further validation in a large-scale study. The glycerophospholipid metabolism pathway can continually provide glycerophospholipids particularly for membrane production of highly proliferating cancer cell [39]. Our funding may support that Glycerophospholipid metabolism as a novel drug target

Table 3 Unconditional logistic regression analysis of the association between differential metabolites and the risk of oral cancer

Metabolites	Univariate Logistic regression		Multivariate Logistic regression	
	OR(95%CI)	*P*	OR(95%CI)*	*P**
PC(18:3e/17:2)	0.66(0.49–0.89)	0.006	0.68(0.49–0.95)	0.024
ACar(16:0)	0.25(0.14–0.44)	< 0.001	0.29(0.17–0.51)	0.011
PC(17:0/20:3)	0.04(0.00-0.37)	0.005	0.03(0.00-0.33)	0.005
TAG(16:0/20:4/20:4)	0.02(0.00-0.20)	0.001	0.01(0.00-0.16)	0.001
ACar(18:2)	0.31(0.16–0.57)	< 0.001	0.34(0.19–0.62)	0.008
PC(16:2/26:4)	0.00(0.00-0.10)	0.002	0.00(0.00-0.03)	0.001
PEtOH(24:4/22:6)	0.00(0.00-0.11)	0.002	0.00(0.00-0.12)	0.002
PC(20:4/20:5)	0.03(0.00-0.48)	0.013	0.02(0.00-0.36)	0.009
TAG(13:0/13:0/21:2)	0.01(0.00-0.26)	0.007	0.00(0.00-0.16)	0.005
SHexCer(d34:1)	5.81(0.89–38.20)	0.067	3.57(0.44–28.78)	0.232
PE(18:5e/21:2)	10.49(3.14–35.06)	< 0.001	10.96(2.76–43.45)	0.001
PC(22:4e/22:6)	15.85(4.44–56.67)	< 0.001	15.12(3.78–60.47)	< 0.001
GlcADG(14:1/14:1)	6.89(2.68–17.69)	< 0.001	6.45(2.45-17.00)	< 0.001

*Adjustment for hypertension and diabetes

against oral cancer in future. A cell-based experiment verified that high levels of Alpha-Linolenic Acid promote apoptosis by activating the JNK/FasL/caspase 8/caspase 3-extrinsic pathway and the Bid/cytochrome c/caspase 9 pathway in OSCC patients [40]. The levels of oleic acid and linoleic acid were lower in 27 OSCC tissues compared with their adjacent normal-appearing tissues [41]. Those may partially support our findings about the Linoleic acid metabolism and alpha-Linolenic acid metabolism changes in oral cancer. Additionally, we found that Glycerolipid metabolism also plays an important role in oral cancer. On the one hand, cancer cells heavily depend on glycerolipids to produce new membranes for their rapid proliferation and division [42]. On the other hand, metabolites derived from Glycerolipid pathways (such as diacylglycerol) can activate several signaling proteins that play major roles in activating cancer-promoting pathways [43]. Previous research has highlighted that the Glycerolipid metabolism pathway is excessively active in patients with head and neck cancers [44]. Interestingly, similar metabolic shifts have also been identified in esophageal squamous cell carcinoma, underscoring a potential commonality in metabolic reprogramming across these cancer types [45]. Future studies should explore the precise mechanisms and causal relationships between those metabolism changes and oral cancer.

However, the limitations should not be ignored. On the one hand, our one-hospital case-control study provides relatively weak causal inference, and further prospective evidence from multiple centers is needed. On the other hand, while in our study, the lipid-metabolite-based disease classification model could not deliver clinically relevant results for the early diagnosis of OC, we believe it could exhibit clinical utility after integrating more clinical data and confirming its validation in larger cohorts.

Conclusion

In conclusion, the glycerophospholipid metabolism pathway exhibited most significant changes in OC patients. The three most significant changes in lipid metabolites, including PC(18:3e/17:2), ACar(14:2), and GlcADG(14:1/14:1), identified in this study show potential as candidate biomarkers for OC diagnosis. A KNN model constructed using 13 metabolites demonstrated excellent performance in distinguish the OC patients and healthy controls.

Abbreviations

OC	Oral cancer
UHPLC-QE-MS	Ultra-high Performance Liquid Chromatography Q Exactive Mass Spectrometry
LASSO	Least Absolute Shrinkage and Selection Operator
RFE	Recursive Feature Elimination
KNN	K-Nearest Neighbor
PCA	Principal Component Analysis
OPLS-DA	The Orthogonal Projections to Latent Structures-Discriminant Analysis
KEGG	The Kyoto Encyclopedia of Genes and Genomes
SVM	Support Vector Machine
RF	Random forest

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-13561-x.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

This study is a joint effort of many investigators and staff members, and their contribution is gratefully acknowledged. We especially thank all patients who participated in this study.

Author contributions

NW, YJC, JLL, and YLL: Formal analysis; writing-original draft; writing-review and editing. FC, FQL, and JW: Methodology; writing-review and editing. HYS, LLS,

WHH, JYH, YQ, LL, LSL and BS: Investigation. LZP and BCH: Conceptualization; funding acquisition; formal analysis; writing-review and editing.

Funding
This study was supported by Fujian Natural Science Foundation Program No, Grant/Award Numbers: 2022J01235 and 2022J01239; Joint Funds for the Innovation of Science and Technology of Fujian Province, Grant/Award Number: 2021Y9015; Hainan Provincial Natural Science Foundation of China (Grant No. 821RC1066).

Data availability
The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Declarations

Ethics approval and consent to participate
Written informed consent was obtained from all enrolled participants. This research was performed in accordance with the ethical standards of the Helsinki Declaration, and the ethical approval was obtained from the Ethics Committees of Fujian Medical University, Fuzhou, China (approval ID: 2011053).

Consent for publication
Not Applicable.

Competing interests
The authors declare no competing interests.

Received: 15 April 2024 / Accepted: 20 January 2025
Published online: 30 January 2025

References

1. Bawaskar HS, Bawaskar PH. Oral diseases: a global public health challenge. Lancet. 2020;395(10219):185–6.
2. Sabharwal HV, Gupta SK, Sharma S, Singh RK, Chhabra KG. The knowledge, perception and behavior among dental practitioners towards diagnosis of oral pathological lesions by biopsy-A cross-sectional survey. J Med Pharm Allied Sci. 2021;10(5):3500–3.
3. Lohe VK, Degwekar SS, Bhowate RR, Kadu RP, Dangore SB. Evaluation of correlation of serum lipid profile in patients with oral cancer and precancer and its association with tobacco abuse. J oral Pathol Med. 2010;39(2):141–8.
4. Kumar P, Augustine J, Urs AB, Arora S, Gupta S, Mohanty VR. Serum lipid profile in oral cancer and leukoplakia: correlation with tobacco abuse and histological grading. J Cancer Res Ther. 2012;8(3):384–8.
5. Xu C, Gu L, Kuerbanjiang M, Jiang C, Hu L, Liu Y, Xue H, Li J, Zhang Z, Xu Q. Adaptive activation of EFNB2/EPHB4 axis promotes post-metastatic growth of colorectal cancer liver metastases by LDLR-mediated cholesterol uptake. Oncogene. 2023;42(2):99–112.
6. Al-Saffar N, Titley J, Robertson D, Clarke P, Jackson L, Leach M, Ronen S. Apoptosis is associated with triacylglycerol accumulation in Jurkat T-cells. Br J Cancer. 2002;86(6):963–70.
7. Bandu R, Mok HJ, Kim KP. Phospholipids as cancer biomarkers: mass spectrometry-based analysis. Mass Spectrom Rev. 2018;37(2):107–38.
8. Santos CR, Schulze A. Lipid metabolism in cancer. FEBS J. 2012;279(15):2610–23.
9. Bian X, Liu R, Meng Y, Xing D, Xu D, Lu Z. Lipid metabolism and cancer. J Exp Med 2021, 218(1).
10. Zhan X, Li J, Guo Y, Golubnitschaja O. Mass spectrometry analysis of human tear fluid biomarkers specific for ocular and systemic diseases in the context of 3P medicine. EPMA J. 2021;12:449–75.
11. McNeish DM. Using lasso for predictor selection and to assuage overfitting: a method long overlooked in behavioral sciences. Multivar Behav Res. 2015;50(5):471–84.
12. Ding Y, Wilkins D. Improving the performance of SVM-RFE to select genes in microarray data. BMC bioinformatics: 2006. Springer; 2006. pp. 1–8.
13. Kamsali S, Swathi B, Rudrakumar M. Predictive Analytics in Healthcare by Leveraging Feature Engineering and Machine Learning.
14. Muhammad G, Naveed S, Nadeem L, Mahmood T, Khan AR, Amin Y, Bahaj SAO. Enhancing prognosis accuracy for ischemic cardiovascular disease using K nearest neighbor algorithm: a robust approach. IEEE Access 2023.
15. Aljameel SS, Khan IU, Aslam N, Aljabri M, Alsulmi ES. Machine learning-based model to predict the Disease Severity and Outcome in COVID-19 patients. Sci Program. 2021;2021(1):5587188.
16. Du Z, Yang Y, Zheng J, Li Q, Lin D, Li Y, Fan J, Cheng W, Chen X-H, Cai Y. Accurate prediction of coronary heart disease for patients with hypertension from electronic health records with big data and machine-learning methods: model development and performance evaluation. JMIR Med Inf. 2020;8(7):e17257.
17. Yang XG, Yang SS, Bao Y, Wang QY, Peng Z, Lu S. Novel machine-learning prediction tools for overall survival of patients with chondrosarcoma: based on recursive partitioning analysis. Cancer Med-Us. 2024;13(15):e70058.
18. Xie X, Luo C, Wu S, Qiao W, Deng W, Jin L, Lu J, Bu L, Duffau H, Zhang J. Recursive partitioning analysis for survival stratification and early imaging prediction of molecular biomarker in glioma patients. BMC Cancer. 2024;24(1):818.
19. Radoï L, Paget-Bailly S, Menvielle G, Cyr D, Schmaus A, Carton M, Guida F, Cénée S, Sanchez M, Guizard A-V. Tea and coffee consumption and risk of oral cavity cancer: results of a large population-based case-control study, the ICARE study. Cancer Epidemiol. 2013;37(3):284–9.
20. Butler LM, Perone Y, Dehairs J, Lupien LE, de Laat V, Talebi A, Loda M, Kinlaw WB, Swinnen JV. Lipids and cancer: emerging roles in pathogenesis, diagnosis and therapeutic intervention. Adv Drug Deliv Rev. 2020;159:245–93.
21. Hoy AJ, Nagarajan SR, Butler LM. Tumour fatty acid metabolism in the context of therapy resistance and obesity. Nat Rev Cancer. 2021;21(12):753–66.
22. Matsushita Y, Nakagawa H, Koike K. Lipid metabolism in oncology: why it matters, how to research, and how to treat. Cancers. 2021;13(3):474.
23. Molendijk J, Kolka CM, Cairns H, Brosda S, Mohamed A, Shah AK, Brown I, Hodson MP, Hennessy T, Liu G. Elevation of fatty acid desaturase 2 in esophageal adenocarcinoma increases polyunsaturated lipids and may exacerbate bile acid-induced DNA damage. Clin Translational Med. 2022;12(5):e810.
24. Wojakowska A, Cole LM, Chekan M, Bednarczyk K, Maksymiak M, Oczko-Wojciechowska M, Jarząb B, Clench MR, Polańska J, Pietrowska M. Discrimination of papillary thyroid cancer from non-cancerous thyroid tissue based on lipid profiling by mass spectrometry imaging. Endokrynologia Polska. 2018;69(1):2–8.
25. Gibellini F, Smith TK. The Kennedy pathway—de novo synthesis of phosphatidylethanolamine and phosphatidylcholine. IUBMB Life. 2010;62(6):414–28.
26. Arlauckas SP, Popov AV, Delikatny EJ. Choline kinase alpha—putting the ChoK-hold on tumor metabolism. Prog Lipid Res. 2016;63:28–40.
27. Iorio E, Ricci A, Bagnoli M, Pisanu ME, Castellano G, Di Vito M, Venturini E, Glunde K, Bhujwalla ZM, Mezzanzanica D. Activation of phosphatidylcholine cycle enzymes in human epithelial ovarian cancer cells. Cancer Res. 2010;70(5):2126–35.
28. Indiveri C, Iacobazzi V, Tonazzi A, Giangregorio N, Infantino V, Convertini P, Console L, Palmieri F. The mitochondrial carnitine/acylcarnitine carrier: function, structure and physiopathology. Mol Aspects Med. 2011;32(4–6):223–33.
29. Ganti S, Taylor SL, Kim K, Hoppel CL, Guo L, Yang J, Evans C, Weiss RH. Urinary acylcarnitines are altered in human kidney cancer. Int J Cancer. 2012;130(12):2791–800.
30. Wenzel U, Nickel A, Daniel H. Increased carnitine-dependent fatty acid uptake into mitochondria of human colon cancer cells induces apoptosis. J Nutr. 2005;135(6):1510–4.
31. Zha S, Ferdinandusse S, Hicks JL, Denis S, Dunn TA, Wanders RJ, Luo J, De Marzo AM, Isaacs WB. Peroxisomal branched chain fatty acid β-oxidation pathway is upregulated in prostate cancer. Prostate. 2005;63(4):316–23.
32. Ni J, Xu L, Li W, Zheng C, Wu L. Targeted metabolomics for serum amino acids and acylcarnitines in patients with lung cancer. Experimental Therapeutic Med. 2019;18(1):188–98.
33. Sachan DS, Dodson WL. The serum carnitine status of cancer patients. J Am Coll Nutr. 1987;6(2):145–50.
34. Van Kaer L, Joyce S. Innate immunity: NKT cells in the spotlight. Curr Biol. 2005;15(11):R429–31.
35. Hill MM, Hemmings BA. Inhibition of protein kinase B/Akt: implications for cancer therapy. Pharmacol Ther. 2002;93(2–3):243–51.
36. Yang T, Hui R, Nouws J, Sauler M, Zeng T, Wu Q. Untargeted metabolomics analysis of esophageal squamous cell cancer progression. J Translational Med. 2022;20(1):127.
37. Zeng W, Zheng W, Hu S, Zhang J, Zhang W, Xu J, Yu D, Peng J, Zhang L, Gong M. Application of lipidomics for assessing tissue lipid profiles

of patients with squamous cell carcinoma. Technol Cancer Res Treat. 2021;20:15330338211049903.
38. Dickinson A, Saraswat M, Joenväärä S, Agarwal R, Jyllikoski D, Wilkman T, Mäkitie A, Silen S. Mass spectrometry–based lipidomics of oral squamous cell carcinoma tissue reveals aberrant cholesterol and glycerophospholipid metabolism—A pilot study. Translational Oncol. 2020;13(10):100807.
39. Dolce V, Rita Cappello A, Lappano R, Maggiolini M. Glycerophospholipid synthesis as a novel drug target against cancer. Curr Mol Pharmacol. 2011;4(3):167–75.
40. Su C-C, Yu C-C, Shih Y-W, Liu K-L, Chen H-W, Wu C-C, Yang Y-C, Yeh E-L, Li C-C. Protective effect of alpha-linolenic acid on human oral squamous cell carcinoma metastasis and apoptotic cell death. Nutrients. 2023;15(23):4992.
41. Askari M, Darabi M, Zare Mahmudabadi R, Oboodiat M, Fayezi S, Mostakhdemin Hosseini Z, Pirzadeh A. Tissue fatty acid composition and secretory phospholipase-A2 activity in oral squamous cell carcinoma. Clin Transl Oncol. 2015;17:378–83.
42. Lim JY, Kwan HY. Roles of lipids in cancer. Adv Lipid Metabolism 2018.
43. Chen X, Liu Q, Chen Y, Wang L, Yang R, Zhang W, Pan X, Zhang S, Chen C, Wu T. Carboxylesterase 2 induces mitochondrial dysfunction via disrupting lipid homeostasis in oral squamous cell carcinoma. Mol Metabolism. 2022;65:101600.
44. Taware R, Taunk K, Pereira JA, Shirolkar A, Soneji D, Câmara JS, Nagarajaram H, Rapole S. Volatilomic insight of head and neck cancer via the effects observed on saliva metabolites. Sci Rep-Uk. 2018;8(1):17725.
45. Chen Y, Yang H, Huang X, Wang R, Mao W, Chen Z. Metabolomics and transcriptomics joint analysis reveals altered amino acid metabolism in esophageal squamous cell carcinoma. 2023.

Publisher's note