Article

Unsupervised Hierarchical Clustering of Head and Neck Cancer Patients by Pre-Treatment Plasma Metabolomics Creates Prognostic Metabolic Subtypes

Ronald C. Eldridge [1,*], Zhaohui S. Qin [2], Nabil F. Saba [3], Madelyn C. Houser [1], D. Neil Hayes [4], Andrew H. Miller [5], Deborah W. Bruner [1], Dean P. Jones [6] and Canhua Xiao [1]

[1] Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, GA 30322, USA
[2] Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA
[3] Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA
[4] Department of Medicine, UT/West Institute for Cancer Research, University of Tennessee Health Science Center, Memphis, TN 38163, USA
[5] Department of Psychiatry and Behavioral Sciences, School of Medicine, Emory University, Atlanta, GA 30322, USA
[6] Division of Pulmonary, Allergy and Critical Care Medicine, Emory University, Atlanta, GA 30322, USA
* Correspondence: ronald.eldridge@emory.edu

Simple Summary: There is a need to identify new translational prognostic biomarkers in head and neck cancer. Metabolomics, the study of small molecules resulting from cellular metabolism, is an emerging and promising field regarding head and neck cancer. We performed metabolomics on patients' blood prior to treatment and found that it can divide patients into high-risk and low-risk groups based on their cancer progression and survival. We believe our study provides compelling results to consider metabolomics as a translational prognostic biomarker and that it may offer novel information for patient risk stratification. With continued research, we hope to gain a fuller understanding of how metabolomics may aid in the early detection, prognosis, treatment monitoring, and targeted therapies of head and neck cancer.

Abstract: There is growing evidence that the metabolism is deeply intertwined with head and neck squamous cell carcinoma (HNSCC) progression and survival but little is known about circulating metabolite patterns and their clinical potential. We performed unsupervised hierarchical clustering of 209 HNSCC patients via pre-treatment plasma metabolomics to identify metabolic subtypes. We annotated the subtypes via pathway enrichment analysis and investigated their association with overall and progression-free survival. We stratified the survival analyses by smoking history. High-resolution metabolomics extracted 186 laboratory-confirmed metabolites. The optimal model created two patient clusters, of subtypes A and B, corresponding to 41% and 59% of the study population, respectively. Fatty acid biosynthesis, acetyl-CoA transport, arginine and proline, as well as the galactose metabolism pathways differentiated the subtypes. Relative to subtype B, subtype A patients experienced significantly worse overall and progression-free survival but only among ever-smokers. The estimated three-year overall survival was 61% for subtype A and 86% for subtype B; log-rank $p = 0.001$. The association with survival was independent of HPV status and other HNSCC risk factors (adjusted hazard ratio = 3.58, 95% CI: 1.46, 8.78). Our findings suggest that a non-invasive metabolomic biomarker would add crucial information to clinical risk stratification and raise translational research questions about testing such a biomarker in clinical trials.

Keywords: metabolomics; head and neck cancer; biomarkers; survival; clustering; fatty acids; smoking; human papillomavirus; arginine and proline; galactose

1. Introduction

Recent treatment advances for head and neck squamous cell carcinoma (HNSCC) have improved patients' survival but at a large cost to their quality of life [1]. The recommended regimen of chemoradiotherapy causes up to 80% of HNSCC patients to experience debilitating side effects such as mucositis, dysphagia, malnutrition, speech problems, lymphedema, and cognitive decline [2–7]. Additionally, more than 50% of these patients remain afflicted years after treatment [8]. This suffering has motivated prognostic biomarker research to help expedite clinical trials of less toxic targeted therapies (e.g., epidermal growth factor receptor inhibitors) and de-escalated standard therapies (e.g., lower doses of radiation) [9].

Human papillomavirus (HPV), and to a lesser extent, tobacco use, are the most widely translated prognostic markers in HNSCC. In 2010, Ang et al. proved that HPV status, when combined with smoking pack-years, could stratify oropharyngeal SCC patients into low, intermediate, and high-risk groups with three-year survivals of 93%, 71%, and 46%, respectively [10]. Their findings spawned a decade of de-escalated therapeutic research to determine if patients deemed low-risk—those who are HPV-positive with a minimal history of tobacco use—could be spared the long-term toxicity of standard treatment. Despite recent setbacks [11,12], this remains an active area of research [13,14]. More recently, circulating HPV DNA has garnered interest as a non-invasive biomarker of post-treatment surveillance, since it can accurately predict tumor recurrence [15]. However, its clinical utility would be limited to HPV-positive oropharyngeal SCC—approximately 30% of the total HNSCC patient population [16].

Other HNSCC prognostic markers being studied include genetic mutations, molecular variations, immunologic signaling, and tumor imaging, but while this is promising, none have been proven to optimize therapy. Functional loss of p53, a tumor suppressor gene, is associated with poor prognosis [17]; however, its high frequency of mutation—upwards of 85% in HPV-unrelated HNSCC [18]—raises questions about its discriminating specificity. Overexpression of the epidermal growth factor receptor (EGFR), a well-studied tyrosine kinase receptor, has shown a prognostic ability, however, similar to the loss of p53, it is also quite common (~90%) in HNSCC [19]. Furthermore, there is no standardized antibody or staining protocol to measure EGFR, and overexpression of EGFR is not consistently associated with a response to EGFR-targeted therapies (e.g., cetuximab) [20], raising additional concerns about its future clinical use. The immunosuppressive nature of the HNSCC microenvironment, relative to cancers at other tissue sites, makes immunologic biomarkers interesting targets. Programmed cell death-1 (PD-1) and its primary ligand (PD-L1) are transmembrane signaling receptors that negatively regulate the adaptive immune response (e.g., cytotoxic T cells). PD-1 and PD-L1 are checkpoint inhibitors being studied as targeted HNSCC immunotherapy after the failure of standard therapy [21,22]. Upregulated PD-L1 is associated with a worse prognosis but more evidence is needed to predict which patients will respond to anti-PD-1 monoclonal antibody immunotherapy [23]. Non-specific immunosuppressive cytokines and chemokines (e.g., IL-10 and TGF-beta) show some ability as prognostic biomarkers of tumor-mediated immunomodulatory signaling but results remain inconsistent [24]. Lastly, cross-sectional tumor imaging via magnetic resonance imaging (MRI), computed tomography (CT), or positron emission tomography/computed tomography (PET/CT), plays an important role in the diagnosis and clinical management of HNSCC. The imaging of HNSCC often provides crucial information for TNM staging [25]. Outside of staging, functional imaging by 18-fluorodexyglucose PET/CT provides a highly sensitive measure of tumor metabolic activity via glucose uptake. Patients whose tumors show higher levels of metabolic activity tend to have a worse prognosis [26], but questions remain about quantifiability and low specificity since non-cancerous tissue and immune cells in the microenvironment also have a high glucose uptake.

Metabolic biomarkers have, in general, been overlooked in HNSCC translational research compared to their genomic, molecular, immunologic, and imaging counterparts. There is, however, growing evidence that the metabolism plays an outsized role in activating the host's immune response [27] and that it is deeply intertwined with HNSCC

progression and survival [28]. In 2020, five independent studies found differential expression of metabolic genes to be prognostic of patient survival [29–33]. While promising, translating a gene-based metabolic biomarker poses some challenges. First, the within-tumor metabolic heterogeneity of HNSCC is extreme, [34] amplifying the probability of sampling bias; moreover, direct sampling of the tumor is costly, invasive, and burdensome on the patient [35,36]. Second, since most metabolic regulation occurs downstream of the genome, [37] a gene-based biomarker is prone to miss vital metabolic transitions and patterns. A non-invasive alternative is to measure circulating metabolites in patients' blood. Recent advances in high-resolution metabolomics allow for comprehensive quantification of the host's metabolism and systemic metabolic signaling. However, as an emerging field, little is known about circulating metabolite patterns in HNSCC patients and their clinical potential. In the following study, we used multivariate unsupervised hierarchical clustering of plasma metabolomics to discover metabolic subtypes from a heterogeneous group of 209 HNSCC patients. We sought to answer three questions: (1) Are there circulating metabolic differences in HNSCC patients? Secondly, (2) if so, are those metabolic differences prognostic of outcomes, namely survival? Lastly, (3) what are the implications for a potential non-invasive metabolomic biomarker in HNSCC?

2. Materials and Methods

2.1. Study Cohort

This was a prospective cohort of 209 HNSCC patients recruited from 2013 to 2016 at radiation oncology clinics affiliated with Emory University Hospital before undergoing intensity-modulated radiation therapy (IMRT) or IMRT with concurrent platinum-based chemotherapy [38]. The inclusion criteria were histological HNSCC with no distant metastasis, ≥21 years of age, and no evidence of uncontrolled metabolic, hematologic, cardiovascular, renal, hepatic, or neurologic disease. Patients with simultaneous primaries or major psychiatric disorders were excluded. Demographic and clinical variables were collected through chart review and standardized questionnaires at study entry. Prior to IMRT, blood was collected into chilled EDTA tubes for the immediate isolation of plasma. Plasma was separated by centrifugation at $1000 \times g$ for 10 min at 4 °C, then aliquoted into siliconized polypropylene tubes and stored at −80 °C.

2.2. Overall and Progression-Free Survival

Vital status and disease progression were ascertained by linking our patient cohort to the Georgia Comprehensive Cancer Registry maintained at Emory University. Overall survival was defined as the time from study entry to death by any cause. Progression-free survival was defined as the time from study entry to HNSCC progression noted in the medical chart or death. Patients who did not experience the outcome were censored at the last date of contact up to June 2022. We were unable to follow 20 subjects and they were excluded from the survival analyses.

2.3. High-Resolution Untargeted Metabolomics (HRM) of Blood Plasma

Our HRM approach used established liquid chromatography-mass spectrometry (LC-MS) protocols developed at the Emory Clinical Biomarkers Laboratory [39,40] (Supplementary File S1). Frozen aliquots were thawed and extracted with ice-cold acetonitrile. Supernatants were added to a 4 °C autosampler in random order. Each sample was divided and analyzed in triplicate (10 µL) using dual chromatography separation: hydrophilic interaction liquid chromatography (HILIC) with positive electrospray ionization and 18-carbon (C18) with negative ionization. Analyte separation was achieved using a 2.1 mm × 100 mm × 2.6 µm Accucore column (Thermo Scientific, Norcross, GA, USA) and a gradient elution of 2% of formic acid, water, and acetonitrile starting at 10%, 10%, and 80%, for 1.5 min with a linear increase to 10%, 80%, and 10% at 6 min and held for 4 min per injection.

The high-resolution Fourier transform Orbitrap mass spectrometer (Dionex Ultimate 3000, Q-Exactive HF, Thermo Scientific) operated in full scan mode at 120,000 resolution and a mass-to-charge ratio (m/z) range of 85.0000–1275.0000. Probe temperature, capillary temperature, sweep gas, and S-Lens RF levels were maintained at 200 °C, 300 °C, 1 arbitrary unit (AU), and 45 AU, respectively. Positive tune settings for sheath gas, auxiliary gas, sweep gas, and spray voltage settings were 45 AU, 25 AU, and 3.5 kV, respectively. Raw data were extracted and aligned using apLCMS [41] and xMSanalyzer [42] separately for the HILIC-positive and C18-negative chromatography runs. Mass, retention time, and ion intensity were measured for each uniquely identified m/z spectral feature. The m/z features with a median coefficient of variation (CV) within technical replicates ≥75% were removed. The remaining feature intensities were median-summarized across sample triplicates with the requirement that ≤1 replicate was imputed. Correction for batch effects was performed using ComBat version: 3.38 [43]. Features missing in <20% of the samples were imputed using k-nearest neighbors and features missing in >20% were removed; we imputed less than 3% of the data points. The remaining features were log2-transformed and Z-score standardized (i.e., subtracting the metabolite-specific population mean and dividing by the metabolite-specific standard deviation). This was done to normalize the data and pool the metabolites across the HILIC-positive and C18-negative chromatography runs. We restricted our statistical investigation to the features that matched m/z within 10 ppm and retention time within 30 s of metabolites that the Emory CBL previously identified using accurate mass MS1 signal, coelution with authentic standard, and ion dissociation spectra (MS^2/MS^n) matching the authentic standard [44]. Meaning, we believe each analyzed metabolite met Level 1 identification according to the Metabolomics Standards Initiative [45] and Schymanski et al. [46]. This left 125 HILIC-positive and 61 C18-negative metabolites to be analyzed.

2.4. Statistical Analysis

Here, we summarize the steps we took to identify prognostic metabolic subtypes in the HNSCC patient population. First, we identified metabolic clusters of HNSCC patients using unsupervised hierarchical clustering of plasma metabolites, an approach that groups HNSCC patients according to metabolite patterns agnostic to patient demographics, clinical factors, or survival. Second, we determined the relative importance of each metabolite to the clustering via post-hoc Wilcoxon rank sum tests. We inputted the meaningful metabolites into pathway enrichment analysis to metabolically characterize the clusters. Third, we examined if the metabolic subtypes were associated with baseline clinical factors (e.g., age, sex, HPV status) and immune markers. Lastly, we modeled the metabolic subtypes with overall and progression-free survival to determine their potential clinical utility.

The unsupervised hierarchical clustering used 186 lab-confirmed plasma metabolites across 209 HNSCC patients. The patients were clustered using JASP version 0.14 according to Euclidean distance and Ward.D linkage, two common algorithmic settings. We fit models allowing for 2–6 patient clusters and chose the optimal model from a variety of clustering metrics: the lowest Bayesian Information Criterion, the highest silhouette scores, the highest Dunn index, the highest Calinski–Harabasz index, and the lowest model entropy. After determining the optimal model, we performed post-hoc Wilcoxon rank sum tests to rank each metabolite according to differential expression between the clusters.

We characterized each cluster via metabolic pathway enrichment analysis using roughly the top 20% of differentially expressed metabolites (i.e., 38 metabolites that met a Benjamini–Hochberg multiple testing adjusted *p*-value < 0.00001). Since we used differential expression as a post-hoc ranking tool, the top 20% cutoff is arbitrary. By limiting the pathway analysis to the top 20% we can describe enriched pathways among the most influential metabolites to the clustering. Changing this cutoff by making it higher or lower may lead to different pathway enrichment findings. Pathway enrichment was performed by Metabolanalyst using the Small Molecule Pathway Database (SMPDB) [47]. We calculated enrichment by comparing the number of pathway hits to the subset of differentially

expressed metabolites relative to the expected number of pathway hits if 38 metabolites were drawn at random from the 186 total using a chi-squared test statistic.

We tested the differences in proportions or means of many clinical factors across the subtypes using chi-square or *t*-tests as appropriate. We also examined whether the metabolic subtypes were associated with patient overall survival via Kaplan–Meier curves and Cox models adjusted for age, sex, race, body mass index, marriage status, HPV status, smoking history, tumor site and stage, treatment, Eastern Cooperative Oncology Group (ECOG) performance, feeding tube, patient-reported prior comorbidities, and circulating levels of albumin, hemoglobin, neutrophil-to-lymphocyte ratio (NLR) and platelet-to-lymphocyte ratio (PLR). We tested the proportional hazards assumption by modeling a subtype-with-time interaction term. We stratified our survival analyses and Kaplan–Meier curves by whether subjects were never vs. ever smokers, then additionally by HPV status (related vs. unrelated). We classified each subject as either alive or dead at three years of follow-up, ran a logistic regression, and plotted a receiver operating characteristic (ROC) curve of three-year survival against three models: an HPV status-only model, a metabolic subtype-only model, and a model with HPV status and metabolic subtypes. Unless stated, all analyses were conducted using R version 4.2.2 [48] and RStudio build 386 [49].

2.5. Sensitivity Analysis

To examine the robustness of our hierarchical clustering approach, we performed two alternative clustering methods, K-means and random forest, and report those cluster groupings with overall survival.

3. Results

Our study population compares favorably to HNSSC patients across the U.S. as estimated by Surveillance, Epidemiology, and End Results (SEER) [50]. In our study, the median age at diagnosis was 60 years old compared to 64 in SEER. A total of 75% of our cases were men and 81% identified as white compared to 71% and 75%, respectively, in SEER. SEER estimates that 67% of HNSCC patients in the U.S. are diagnosed with regional lymph node or distant metastasis, and 68% of patients live longer than five years; our proportions were slightly higher at 74% and 77%, respectively. In our data, the average follow-up time for the subjects who died was 2.4 years compared to 5.3 years for those who were censored; only 8% of the censored subjects had fewer than 3 years of follow-up. Approximately 61% of our patients reported a current or former history of smoking, and 48% were HPV-related. Half of the tumors occurred in the oropharynx. Across many demographic and clinical characteristics, only age statistically differed between the subtypes (Table 1 and Supplementary Figure S1).

A total of 186 lab-confirmed metabolites were used to cluster 209 HNSCC patients: 9 amines; 59 amino acids or amino acid-derivatives; 9 carbohydrates; 17 cholines, betaines, carnitines, or Co-As; 31 lipids; 7 nucleic acids, nucleosides, or nucleotides; 31 simple organic acid derivatives; 11 vitamins or cofactors; and 12 metabolites that do not fit those categories (Supplementary Table S1). The metabolites covered 88 different human metabolic pathways per the SPMDB reference database (Supplementary Table S2). Furthermore, the median pair-wise Pearson correlation coefficient between metabolites was 0.04 with an inter-quartile range from -0.03 to 0.12 (Supplementary Figure S2), suggesting robust metabolite heterogeneity. Despite this heterogeneity, certain groups of metabolites, namely fatty acids, were highly correlated (Supplementary Figure S3 and Supplementary Table S3).

The optimal model, according to five of six clustering metrics (Supplementary Table S4), created two patient clusters, herein named metabolic subtypes A and B (Figure 1). We refrain from further labeling the subtypes dues to the relative, not absolute, nature of the metabolic differences. Subtypes A and B corresponded to 41% and 59% of our HNSCC study population, respectively. In Wilcoxon rank sum tests, 38 of the 186 metabolites were differentially expressed at a multiple testing adjusted *p*-value < 0.00001 (Supplementary Table S1). From those 38, the top enriched metabolic pathways were fatty acid biosynthesis

(four metabolites; $p = 0.004$), transfer of acetyl groups into mitochondria (three metabolites; $p = 0.03$), arginine and proline metabolism (five metabolites; $p = 0.06$), and galactose metabolism (three metabolites; $p = 0.07$) (Supplementary Table S2). To further explain the enrichment calculation, among our 186 total metabolites, six mapped to the fatty acid biosynthesis pathway. Among those six, four were ranked in the top 38, representing a relatively rare event ($\chi2 = 8.15$, df = 1, $p = 0.004$) if we drew 38 metabolites from the 186 at random. The metabolite Z-scores suggest relatively higher levels of biosynthesis pathways for fatty acids and amino acids, and glycolytic energy metabolism was found in samples from patients in subtype A versus subtype B (Table 2).

Table 1. Descriptive table of 209 HNSCC patients stratified by metabolic subtypes A and B derived from hierarchical clustering of 186 plasma metabolites.

		All (N = 209)	Subtype A (N = 86)	Subtype B (N = 123)	*p* Value
Age	mean (SD)	59.3 (10.1)	61.7 (10.3)	57.6 (9.7)	0.004
BMI	mean (SD)	27.5 (5.2)	28.1 (5.8)	27.1 (4.7)	0.18
Sex	Male	156 (75%)	65 (76%)	91 (74%)	0.92
	Female	53 (25%)	21 (24%)	32 (26%)	
Race	White	170 (81%)	73 (85%)	97 (79%)	0.36
	Black	39 (19%)	13 (15%)	26 (21%)	
HPV status	Negative	108 (52%)	43 (50%)	65 (53%)	0.79
	Positive	101 (48%)	43 (50%)	58 (47%)	
Smoking history	Never	81 (39%)	33 (39%)	48 (39%)	0.83
	Former	67 (33%)	29 (35%)	38 (31%)	
	Current	58 (28%)	22 (26%)	36 (30%)	
Alcohol	<1 drink/week	114 (55%)	47 (55%)	67 (55%)	0.99
	1+ drink/week	92 (45%)	38 (45%)	54 (45%)	
Marital status	Married or partnered	148 (71%)	62 (72%)	86 (70%)	0.85
	Single	61 (29%)	24 (28%)	37 (30%)	
Tumor site	Oropharynx	106 (51%)	45 (53%)	61 (50%)	0.56
	Oral cavity	31 (15%)	12 (14%)	19 (15%)	
	Larynx	35 (17%)	11 (13%)	24 (20%)	
	Other	36 (17%)	17 (20%)	19 (15%)	
Tumor stage *	I	10 (5%)	5 (6%)	5 (4%)	0.77
	II	16 (8%)	5 (6%)	11 (9%)	
	III	83 (40%)	36 (42%)	47 (39%)	
	IV	98 (47%)	40 (47%)	58 (48%)	
T	1	37 (18%)	19 (22%)	18 (15%)	0.43
	2	51 (25%)	20 (24%)	31 (26%)	
	3	47 (23%)	21 (25%)	26 (22%)	
	4	70 (34%)	25 (29%)	45 (37%)	
N	0	46 (22%)	18 (21%)	28 (23%)	0.09
	1	22 (11%)	4 (5%)	18 (15%)	
	2	130 (63%)	60 (70%)	70 (58%)	
	3	8 (4%)	4 (5%)	4 (3%)	
Treatment	Radiotherapy	45 (22%)	20 (23%)	25 (20%)	0.87
	Chemoradiotherapy with Cisplatin	118 (56%)	48 (56%)	70 (57%)	
	Chemoradiotherapy with Carboplatin and Paclitaxel	46 (22%)	18 (21%)	28 (23%)	
Feeding tube	No	80 (40%)	35 (43%)	45 (39%)	0.60
	Yes	118 (60%)	46 (57%)	72 (62%)	
ECOG performance	Active	100 (50%)	39 (48%)	61 (52%)	0.89
	Restricted	73 (37%)	31 (38%)	42 (36%)	
	Non-working	26 (13%)	11 (14%)	15 (13%)	
Prior comorbidity	Yes	154 (75%)	66 (77%)	88 (73%)	0.69
	No	52 (25%)	20 (23%)	32 (27%)	
Albumin	mean (SD)	3.95 (0.42)	3.97 (0.39)	3.94 (0.43)	0.60
Hemoglobin	mean (SD)	13.18 (1.80)	13.22 (1.71)	13.15 (1.87)	0.80
Neutrophil-to-lymphocyte ratio	mean (SD)	3.18 (2.07)	3.32 (1.98)	3.08 (2.13)	0.45
Platelet-to-lymphocyte ratio	mean (SD)	170,347 (98,859)	161,891 (64,755)	176,207 (116,774)	0.35

HNSCC = head and neck squamous cell carcinoma, BMI = body mass index, HPV = human papillomavirus, ECOG = Eastern Cooperative Oncology Group, SD = standard deviation. * Using the 8th edition of the American Joint Committee on Cancer Tumor Staging by Site.
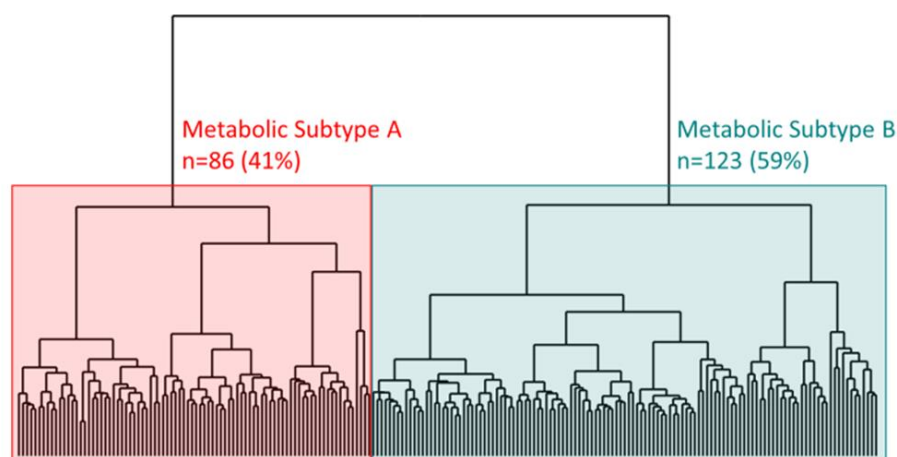
Figure 1. Dendrogram showing two metabolic subtypes, A (red) and B (teal), identified from an unsupervised hierarchical clustering analysis of 209 head and neck cancer patients using 186 lab-confirmed plasma metabolites. Subtype A, the high-risk type showing relatively higher amounts of fatty acid, acetyl Co-A, and amino acid biosynthesis as well as glycolytic metabolism, represents 41% of the sample population and subtype B, the remaining 59%.

Table 2. Fourteen lab-confirmed metabolites that comprised the enriched metabolic pathways that differentiated subtypes A and B.

Name	mz	rt	ESI	Adduct	HMDB#	Z-Score * A	Z-Score * B	*p*-Value **
Fatty acid biosynthesis (*p* = 0.004)
Acetoacetate	101.0244	25	C18-	(M − H)	HMDB0000060	0.53	−0.37	2.0×10^{-9}
β-hydroxybutyrate	103.0401	22	C18-	(M − H)	HMDB0000357	0.43	−0.30	4.3×10^{-6}
FA 16:0 (Palmitate)	255.2329	231	C18-	(M − H)	HMDB00220	0.38	−0.27	4.6×10^{-6}
FA 14:0 (Myristate)	227.2016	212	C18-	(M − H)	HMDB00806	0.38	−0.27	7.1×10^{-6}
Transfer of acetyl groups into the mitochondria (*p* = 0.03)								
Glucose	215.0328	21	C18-	(M + Cl)	HMDB0000122	0.40	−0.28	3.1×10^{-6}
Citric acid	191.0197	19	C18-	(M − H)	HMDB0000094	0.41	−0.28	7.5×10^{-6}
Malic acid	133.0143	20	C18-	(M − H)	HMDB0000156	0.37	−0.26	7.6×10^{-6}
Arginine and Proline metabolism (*p* = 0.06)								
S-adenosylmethionine	399.1445	162	HILIC+	(M + H)	HMDB0001185	0.47	−0.33	1.6×10^{-9}
Proline	116.0706	87	HILIC+	(M + H)	HMDB0000162	0.49	−0.35	8.3×10^{-8}
Ornithine	133.0972	125	HILIC+	(M + H)	HMDB0000214	0.45	−0.32	5.1×10^{-7}
Citrulline	176.103	109	HILIC+	(M + H)	HMDB0000904	0.40	−0.28	8.4×10^{-7}
Guanidinoacetate	118.0617	89	HILIC+	(M + H)	HMDB0000128	0.41	−0.29	8.8×10^{-7}
Galactose metabolism (*p* = 0.07)								
Fructose	219.0265	73	HILIC+	(M + K)	HMDB0000660	0.60	−0.42	3.6×10^{-11}
Galactose	203.0526	71	HILIC+	(M + Na)	HMDB0000143	0.59	−0.41	3.6×10^{-11}
Glucose	215.0328	21	C18-	(M + Cl)	HMDB0000122	0.40	−0.28	3.1×10^{-6}

The mz = mass-to-charge ratio, rt = retention time, ESI = electrospray ionization method, FA = fatty acid, HMDB = human metabolome database. * Z-scores were calculated by subtracting the metabolite-specific population mean and dividing by the standard deviation. ** Wilcoxon rank sum *p*-values adjusted for multiple testing via the Benjamini–Hochberg method.

Relative to subtype B, subtype A patients experienced significantly worse overall and progression-free survival. The estimated 3-year overall survival was 73.3% for subtype A patients and 88.3% for subtype B; log-rank p = 0.003 (Figure 2a). The estimated 3-year progression-free survival was 68.2% for subtype A patients and 79.4% for subtype B; log-rank p = 0.05 (Supplementary Figure S4a). However, the survival pattern varied by smoking history. Among never smokers, overall and progression-free survival did not differ by subtype—the 3-year overall survival was 93.1% for subtype A patients and 91.3% for subtype B; log-rank p = 0.46 (Figures 2b and S4b). In contrast, among ever smokers, subtype A had significantly worse survival compared to subtype B—the 3-year overall survival was 61.0% for subtype A patients and 86.3% for subtype B; log-rank p = 0.001 (Figures 2c and S4c). Further stratifying the ever smoker population by HPV status showed

that the metabolic subtype remained significantly associated with overall and progression-free survival regardless of HPV status (Figures 2d and S4d). Among HPV-unrelated smokers, the estimated 3-year overall survival was 50.7% for subtype A patients and 80.9% for subtype B (*p* = 0.03). Among HPV-related smokers, the estimated 3-year overall survival was 77.4% for subtype A patients and 95.8% for subtype B (*p* = 0.008). Thus, the association with survival was dependent on smoking history but indifferent to HPV status. Intriguingly, despite the importance of having a smoking history, the smoking-related metabolites nicotine (Wilcoxon rank sum *p* = 0.94) and cotinine (Wilcoxon rank sum *p* = 0.06) played almost no role in classifying a patient as subtype A or B (Supplementary Table S1).
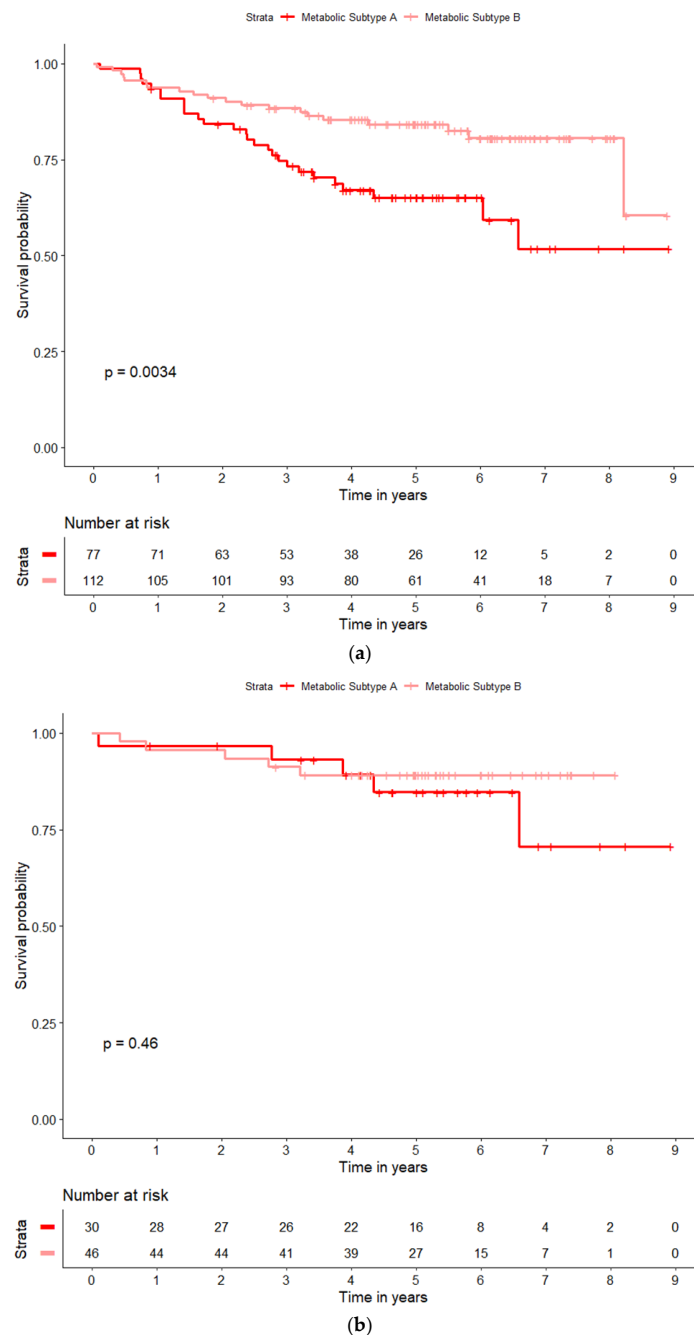

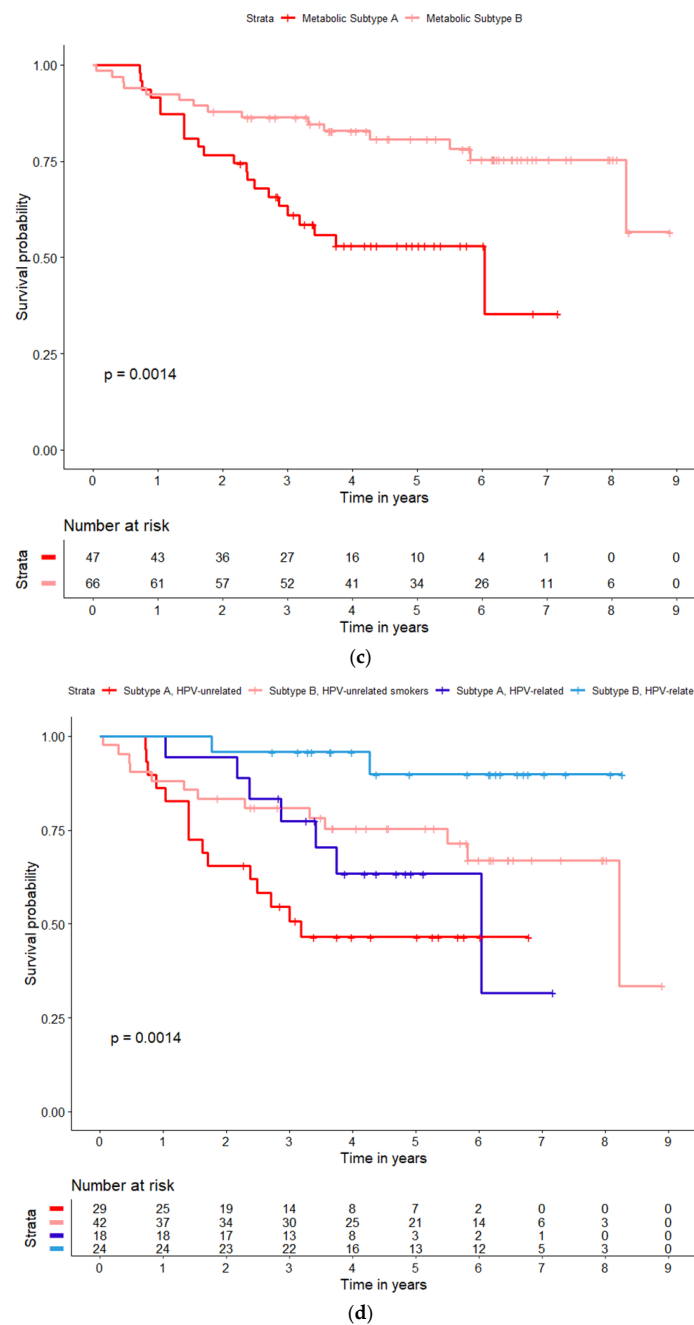(a)

(b)

Figure 2. *Cont.*

(c)

(d)

Figure 2. (**a–d**). Kaplan–Meier overall survival curves by Metabolic Subtypes A and B. (**a**) is among the full population (N = 47 deaths) in which the estimated 3-year survival is 73.3% for Subtype A (dark red) and 88.3% for Subtype B (light red); the curves are statistically significant with a log-rank *p*-value = 0.003. (**b**) shows the metabolic subtype overall survival curves (A vs. B) amongst never smokers (N = 10 deaths). The estimated 3-year survival is 93.1% for subtype A (dark red), and 91.3% for subtype B (light red); *p* = 0.46. (**c**) shows the metabolic subtype overall survival curves (A vs. B) amongst ever smokers (N = 37 deaths). The estimated 3-year survival is 61.0% for subtype A (dark red), and 86.3% for subtype B (light red); *p* = 0.001. (**d**) further stratifies the survival curves by HPV status (unrelated and related) amongst ever smokers (N = 37 deaths). The estimated 3-year survival for the four groups is 50.7% for Subtype A, HPV-unrelated (dark red); 80.9% for Subtype B, HPV-unrelated (light red); 77.4% for subtype A, HPV-related (dark blue); and 95.8% for subtype B, HPV-related (light blue).

The association of the metabolic subtypes with overall survival did not meaningfully change after adjustment by other factors (Table 3). The overall (*p* = 0.58) and progression-

free ($p = 0.37$) cox models met the proportional hazards' assumption testing a subtype by the time interaction term. In age, sex, HPV, and smoking-adjusted Cox models subtype A patients had three times the risk of mortality of subtype B (hazard ratio = 2.98, 95% CI: 1.48, 5.99) among current and former smokers. Further adjustment for race, body mass index, alcohol consumption, marital status, tumor site and stage, treatment, ECOG performance, gastronomy tube, prior comorbidities, and circulating immune markers NLR and PLR only strengthened the point estimates (HR = 3.58, 95% CI: 1.46, 8.78). Given how balanced the risk factors were between the subtypes, as shown in Table 1, it is unsurprising that covariate adjustment did not impact the association between subtype and survival. Similar to the Kaplan–Meier results, we observed no significant associations among never smokers and saw no pattern that would suggest a larger sample size would produce a significant association. The progression-free survival results followed the same pattern as the overall survival results but were slightly attenuated (Table 4). Of note, subtype A had nine times the mortality of subtype B (HR = 9.24, 95% CI: 1.64, 52.1) in age and sex-adjusted models among HPV-related smokers; however, we believe this may be better interpreted as a preliminary finding due to the small number of deaths (N = 9) in this sub-stratum.

Table 3. Estimated associations of metabolic subtypes A vs. B (referent) with patient overall survival via Cox models among the full HNSCC population and stratified by smoking status.

	HR	95% CI	*p* Value
Full Population (N = 189, N_{deaths} = 47)			
Unadjusted model	2.33	(1.30, 4.16)	0.004
Age, sex, HPV, smoking adjusted model	2.38	(1.30, 4.33)	0.005
Fully adjusted model [a]	2.76	(1.32, 5.77)	0.007
Among Ever Smokers (N = 113, N_{deaths} = 37)			
Unadjusted model	2.91	(1.47, 5.78)	0.002
Age, sex, HPV adjusted model	2.98	(1.48, 5.99)	0.002
Fully adjusted model [a]	3.58	(1.46, 8.78)	0.005
Among Never Smokers (N = 76, N_{deaths} = 10)			
Unadjusted model	1.60	(0.46, 5.51)	0.46
Age, sex, HPV adjusted model	1.13	(0.30, 4.24)	0.85
Fully adjusted model [a]	0.92	(0.06, 15.3)	0.95

[a] Further adjusted for race, body mass index, alcohol history, marital status, tumor site and stage, treatment, Eastern Cooperative Oncology Group (ECOG) performance, gastronomy tube, prior comorbidities, albumin, hemoglobin, neutrophil-to-lymphocyte ratio, and platelet-to-lymphocyte ratio; 40 subjects (25 ever, and 15 never smokers), including 5 deaths (4 ever, and 1 never smokers), were removed due to missing covariates. HNSCC = head and neck cancer squamous cell carcinoma, HR = hazard ratio, CI = confidence interval, HPV = human papillomavirus.

Table 4. Estimated associations of metabolic subtypes A vs. B (referent) with progression-free survival via Cox models among the full HNSCC population and stratified by smoking status.

	HR	95% CI	*p* Value
Full Population (N_{events} = 62)			
Unadjusted model	1.62	(0.98, 2.67)	0.06
Age, sex, HPV, smoking adjusted model	1.65	(0.99, 2.75)	0.06
Fully adjusted model [a]	1.70	(0.93, 3.11)	0.08
Among Ever Smokers (N_{events} = 47)			
Unadjusted model	2.06	(1.14, 3.71)	0.02
Age, sex, HPV adjusted model	2.13	(1.17, 3.89)	0.01
Fully adjusted model [a]	2.11	(1.03, 4.32)	0.04
Among Never Smokers (N_{events} = 15)			
Unadjusted model	1.00	(0.35, 2.80)	0.99
Age, sex, HPV adjusted model	0.80	(0.27, 2.37)	0.68
Fully adjusted model [a]	0.48	(0.08, 3.03)	0.44

[a] Further adjusted for race, body mass index, alcohol history, marital status, tumor site and stage, treatment, Eastern Cooperative Oncology Group (ECOG) performance, gastronomy tube, prior comorbidities, albumin, hemoglobin, neutrophil-to-lymphocyte ratio, and platelet-to-lymphocyte ratio; 40 subjects (25 ever and 15 never smokers), including 5 events (4 ever and 2 never smokers), were removed due to missing covariates. HNSCC = head and neck cancer squamous cell carcinoma, HR = hazard ratio, CI = confidence interval, HPV = human papillomavirus.

We used logistic regression and a ROC curve to compare the ability of three models (HPV-only model vs. metabolic subtype-only model vs. HPV-plus metabolic subtype model) to predict 3-year overall survival in HNSCC patients with a history of smoking. The ROC curves in Figure 3 show that, in our data, adding pretreatment metabolic subtypes to the clinically accepted model—a model that uses only HPV—would improve the model's classification of high-risk/low-risk by 20% (AUC of the model with metabolic subtypes = 0.73 vs. AUC of the model with HPV alone = 0.61).
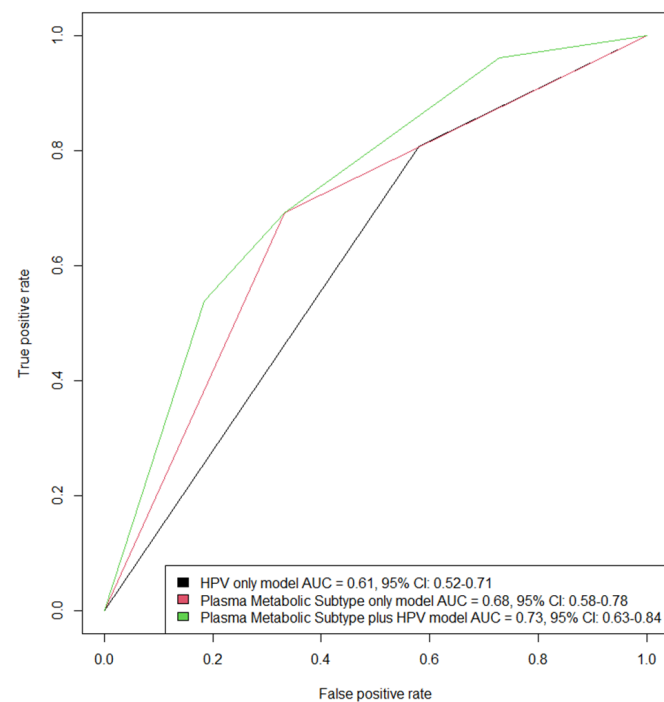
Figure 3. Receiver Operating Characteristic (ROC) curves showing the ability of three logistic regression models to classify death within three years of follow-up for head and neck cancer patients with a prior smoking history. The first model (black line) uses only HPV status. The second model (red line) uses only the pretreatment plasma metabolomic subtype biomarker. The third model (green line) adds the metabolomic subtype to a model of HPV status. Comparing the area under the curve (AUC) of the first and third models indicates a 20% relative increase in predictive ability when the metabolomic subtype biomarker is added to a model of HPV status ($0.73/0.61 = 1.197$).

In the sensitivity analyses (Supplementary Table S5), the random forest clustering showed a similar pattern to the hierarchical clustering, classifying 46% of the HNSCC population into a high-risk group and 54% into a low-risk group. Compared to the low-risk group, the high-risk group had a statistically significant risk of death among HNSCC ever smokers (fully adjusted Cox model HR = 6.75, 95% CI: 2.24, 20.3). The K-means clustering classified 42% into a high-risk group and 58% into a low-risk group. It also found a statistically significant risk of death among HNSCC current and former smokers (fully adjusted Cox model HR = 2.75, 95% CI: 1.09, 6.95).

4. Discussion

Using 186 lab-confirmed plasma metabolites, unsupervised hierarchical clustering parsed 209 HNSCC patients into two metabolic subtypes differentiated by the fatty acid, Co-A, amino acid, and galactose metabolism pathways. The subtypes were associated with a high-risk and low-risk of cancer progression and mortality but only among current and former smokers. Though our results should be replicated, our data suggest that a systemic metabolomic biomarker would be independent of other risk factors, most importantly HPV status, and may add crucial information to risk stratification for either HPV-related or

HPV-unrelated HNSCC patients with a history of smoking. Our findings raise relevant and timely research questions about adding a non-invasive metabolomic blood biomarker as translational research in HNSCC clinical trials.

We believe this is the first study to cluster HNSCC patients by metabolites and investigate survival, complicating the task of comparing our results to the prior literature. A total of eleven HNSCC studies have investigated metabolomics in patient blood: six comparing cancer vs. controls [51–56], four comparing pre- vs. post-treatment [57–60], and one comparing extranodal extension [61], but none investigated survival. Despite the differences in approach, we see similarities between our findings and those of prior research in relation to sugar metabolism (glucose, galactose, fructose), amino acids (proline), acetyl CoA transport (citrate), and fatty acid biosynthesis (acetoacetate, β-hydroxybutyrate). Notably, Li et al.'s metabolic pathway analysis found enrichment in arginine and proline metabolism and CoA biosynthesis [56], matching our enrichment findings. In addition, Rodríguez-Tomàs et al. found advanced prognosis to be associated with pre-treatment plasma β-hydroxybutyrate [58]—an epigenetic signaling metabolite that we found crucial to our clustering.

The enriched metabolic pathways that differentiated the subtypes centered around fatty acid and amino acid biosynthesis and glycolytic metabolism. However, since individual metabolites are often linked to multiple pathways these differences likely represent broader generalities. Our strongest findings were fatty acid biosynthesis and its precursor pathway as well as the transfer of acetyl-CoA to the mitochondria—a vital pathway in the tricarboxylic acid (TCA) cycle. Fatty acid synthesis is an energy storage pathway in which carbons are added to acetyl-CoA via the enzyme fatty acid synthase (FAS) to create longer chain fatty acids necessary for cellular proliferation. In HNSCC, upregulation of FAS and its regulator, epidermal growth factor receptor (EGFR), is positively associated with HPV-negative tumor progression and metastasis [62]. Moreover, FAS and EGFR inhibitors are under active investigation as targets for HNSCC treatment and resistance [63,64]. Likewise, the arginine and proline pathway is central to the biosynthesis of key amino acids for proliferation. Li and colleagues found arginine biosynthesis and other amino acid pathways positively associated with HNSCC compared to controls [55]. Additionally, two arginine and proline pathway metabolites we observed, S-adenosylmethionine and guanidinoacetate, are the metabolic precursors to creatine and phosphocreatine, which are inversely associated with pathological extranodal extension, an adverse prognostic factor in HPV-negative HNSCC [61]. Galactose metabolism, via the Leloir pathway, converts galactose to glucose 1-phosphate to undergo glycolysis, an energy expenditure pathway that is perhaps the most studied metabolic pathway in HNSCC [65]. Wu and colleagues found galactose metabolism positively associated with HNSCC prognosis [29], and in prior research, we found it associated with HPV status and survival [66]. Uptake of glucose is also a prognostic marker in HNSCC imaging [26], and there is considerable interest in targeting or inhibiting glycolysis to improve HNSCC response to radiotherapy [67,68].

An intriguing yet unexplained finding was the prognostic dependency of the subtypes on a prior smoking history. Since the subtypes had similar proportions of current, former, and never smokers, this means their association with survival was dependent on but not confounded by smoking. Unfortunately, the questionnaires implemented by the parent study did not collect data to investigate interaction across pack-years but we note that levels of the smoking-related metabolites nicotine and cotinine did not play a role in subtype clustering and were not associated with survival. We do not yet have a biological rationale for this finding. However, in a somewhat parallel finding, Gutkind et al.'s phase II trial reported a successful histological response to the metabolic diabetes drug metformin only among current and former smokers [69]. Gutkind and colleagues showed that the response was linked to suppression of the mammalian target of the rapamycin (mTOR) pathway, a proliferative pathway that regulates fatty acid biosynthesis [70] and acetyl-CoA [71] but is also regulated by arginine metabolism [72]. There are ongoing and recently completed phase II clinical trials investigating the safety and efficacy of adding metformin to improve

HNSCC patient response to standard chemoradiotherapy or immunotherapy [73,74]. For those trials, and future HNSCC clinical trials of metformin or other mTOR inhibitors, we believe a metabolomic biomarker would be a valuable translational addition.

The prognostic ability of the metabolic subtypes was independent of HPV status, and this may be particularly pertinent to HPV-related HNSCC patients with an extensive smoking history. HPV-related HNSCC smokers are an intermediate-risk group whose prognosis appears too uncertain to test de-escalated treatment [10,75]. Among this group of patients, metabolic subtype A showed a nearly ten-fold change in survival relative to subtype B. Though our analysis does not disentangle why smoking contributes to worse outcomes in HPV-related disease, it does suggest that a metabolomic biomarker might aid clinical risk stratification for this varied patient group.

A specific limitation of the study was the inability to assess disease-induced cachexia, muscle loss, or another abnormal nutritional state that may cause metabolic dysregulation. We did not have information on muscle mass, nor did we have information on weight loss prior to study entry; however, the average body mass index between the metabolic subtypes was similar at baseline and that similarity did not change during the first year of follow-up. Additionally, patients were not instructed to fast before blood collection, and we had limited hematological variables from which to assess anemia or nutrition deficiency; we did not have serum hematocrit, bilirubin, iron, ferritin, lactate dehydrogenase, or aminotransaminases. The estimated prevalence of pretreatment cachexia in HNSCC is upwards of 20% and it is associated with poor prognosis and mortality [76], making it a potential confounder. We attempted to mitigate this confounding by adjusting the survival models by factors correlated with a patient's nutritional state, such as body mass index, albumin and hemoglobin levels, as well as the presence of a gastronomy tube but we cannot rule out residual confounding. Metabolomics may be intimately involved with cancer patient cachexia. In a very small number of HNSCC patients, Boguszewicz et al. showed evidence that increased serum levels of ketone bodies (β-hydroxybutyrate, acetone, and acetoacetate) along with decreased serum glucose may provide a real-time signal of treatment-related toxicity and future cachexia [77]. Though we used pretreatment LC-MS metabolomics and Boguszewicz et al. used posttreatment ^{1}H NMR metabolomics, the similarities between our two findings warrant further investigation.

Other limitations include a relatively small sample size, the lack of external validation, and the inability to link our plasma results with a tumor. Though a sample size of 200 is not particularly small, the stability and reproducibility of unsupervised clustering are improved with larger numbers and multiple datasets. Unfortunately, unlike genomics datasets, metabolomics datasets are scarce and not publicly available. Our results need to be validated, and we are actively working to do so. However, since we were analyzing relative intensity rather than metabolite concentration further standardization would be required before pooling data with another study. The recruitment protocol of the parent study limited our metabolomics analyses to patient plasma. Without paired tumor specimens, we cannot investigate tumor characteristics such as p53 or Ki67 staining. We also cannot directly compare our circulating metabolic subtypes to tumor genomic subtypes found in the prior literature. Moreover, we cannot be sure circulating metabolites adequately represent the metabolism occurring within the tumor and its microenvironment. Therefore, our study is more akin to biomarker discovery from systemic metabolic signals rather than the metabolic classification of HNSCC. However, it is noteworthy that the metabolic pathways we found in patient plasma were also found in tumor genomic subtypes in prior research [78]. Limiting our analysis to lab-confirmed metabolites mitigated issues with metabolite annotation but opens the possibility of missing important metabolic signals. Still, with 186 confirmed metabolites clustering 209 patients, there is a potential issue of high dimensional data making the clusters less reproducible. We explored reproducibility using alternative clustering algorithms, random forest and k means clustering, and we found similar results in both to our primary findings.

5. Conclusions

In a first-of-its-kind investigation, we believe our study provides compelling results to consider metabolomics as a translational prognostic biomarker of HPV-related and HPV-unrelated HNSCC tobacco users. However, there are still pressing research gaps to fill regarding independent validation, integration with other omics data, metabolic interventions, and incorporation into clinical trials before the potential for metabolomics is fully recognized. Addressing these gaps is crucial to gain a fuller understanding of how metabolomics may aid in the early detection, prognosis, treatment monitoring, and targeted therapies of HNSCC.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers15123184/s1, File S1: Standard operating procedure for high-resolution metabolomics. Table S1: List of 186 lab-confirmed metabolites used in the hierarchical clustering. Table S2: list of metabolic pathways tested for enrichment. Table S3: Pearson's correlations of 186 metabolites. Table S4: clustering metrics used to determine the optimal number of patient clusters in the model. Table S5: random Forest and K-means clustering with estimated associations of overall survival. Figure S1: a stratified boxplot of the age distribution by metabolic subtype. Figure S2: a histogram of pair-wise Pearson correlation coefficients between the 186 plasma metabolites used for hierarchical clustering. Figure S3: a clustered heatmap of pair-wise metabolite correlations. Figure S4: progression-free Kaplan–Meier survival curves by metabolic subtypes. References [79–82] are cited in Supplementary Materials.

Author Contributions: Conceptualization: R.C.E., C.X. and D.P.J.; Methodology: R.C.E., Z.S.Q. and D.P.J.; Formal Analysis: R.C.E. and Z.S.Q.; Funding acquisition: R.C.E. and C.X.; Project administration & Data curation: R.C.E. and C.X.; Resources: C.X., D.P.J. and N.F.S.; Software: R.C.E.; Supervision: C.X., D.W.B. and D.P.J.; Writing original draft: R.C.E.; Review & editing manuscript: R.C.E., Z.S.Q., N.F.S., M.C.H., D.N.H., A.H.M, D.W.B., D.P.J. and C.X. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Emory University Institutional Review Board. The parent study (IRB00070167) was originally approved by Emory IRB on 31 October 2013 and has yearly continuing review most recently approved on 15 September 2022. The Emory IRB determined the secondary analysis pertinent to this research was exempt via D4i on 27 June 2019.

Informed Consent Statement: Informed consent was obtained for all subjects involved in this study.

Data Availability Statement: The data are available on request to the corresponding author. De-identified metabolomics data will be deposited at the UC San Diego Metabolomics Workbench. Computer code for the metabolomics analysis using xmsPANDA is freely available at https://github.com/kuppal2/ (accessed on 28 February 2023). Code for the clustering and survival analyses are available upon request to the corresponding author.

Conflicts of Interest: The authors have no conflict of interest related to the work in the manuscript. Unrelated to this work, A.H.M. discloses as a paid consultant for Cerevel Therapeutics and Sirtsei Pharmaceuticals, Inc.

References

1. Winquist, E.; Head and Neck Disease Site Group; Agbassi, C.; Meyers, B.M.; Yoo, J.; Chan, K.K.W. Systemic therapy in the curative treatment of head and neck squamous cell cancer: A systematic review. *J. Otolaryngol.-Head Neck Surg.* **2017**, *46*, 29. [CrossRef] [PubMed]
2. Trotti, A.; Bellm, L.A.; Epstein, J.B.; Frame, D.; Fuchs, H.J.; Gwede, C.K.; Komaroff, E.; Nalysnyk, L.; Zilberberg, M.D. Mucositis incidence, severity and associated outcomes in patients with head and neck cancer receiving radiotherapy with or without chemotherapy: A systematic literature review. *Radiother. Oncol.* **2003**, *66*, 253–262. [CrossRef] [PubMed]
3. Manikantan, K.; Khode, S.; Sayed, S.I.; Roe, J.; Nutting, C.M.; Rhys-Evans, P.; Harrington, K.J.; Kazi, R. Dysphagia in head and neck cancer. *Cancer Treat. Rev.* **2009**, *35*, 724–732. [CrossRef] [PubMed]
4. Deng, J.; Ridner, S.H.; Dietrich, M.S.; Wells, N.; Wallston, K.A.; Sinard, R.J.; Cmelak, A.J.; Murphy, B.A. Prevalence of Secondary Lymphedema in Patients with Head and Neck Cancer. *J. Pain Symptom Manag.* **2012**, *43*, 244–252. [CrossRef] [PubMed]
5. Van Deudekom, F.J.; Schimberg, A.S.; Kallenberg, M.H.; Slingerland, M.; van der Velden, L.-A.; Mooijaart, S.P. Functional and cognitive impairment, social environment, frailty and adverse health outcomes in older patients with head and neck cancer, a systematic review. *Oral Oncol.* **2017**, *64*, 27–36. [CrossRef] [PubMed]
6. Capuano, G.; Gentile, P.C.; Bianciardi, F.; Tosti, M.; Palladino, A.; Di Palma, M. Prevalence and influence of malnutrition on quality of life and performance status in patients with locally advanced head and neck cancer before treatment. *Support. Care Cancer* **2010**, *18*, 433–437. [CrossRef]
7. Jacobi, I.; van der Molen, L.; Huiskens, H.; van Rossum, M.A.; Hilgers, F.J.M. Voice and speech outcomes of chemoradiation for advanced head and neck cancer: A systematic review. *Eur. Arch. Oto-Rhino-Laryngol.* **2010**, *267*, 1495–1505. [CrossRef]
8. Denis, F.; Garaud, P.; Bardet, E.; Alfonsi, M.; Sire, C.; Germain, T.; Bergerot, P.; Rhein, B.; Tortochaux, J.; Calais, G. Final Results of the 94–01 French Head and Neck Oncology and Radiotherapy Group Randomized Trial Comparing Radiotherapy Alone with Concomitant Radiochemotherapy in Advanced-Stage Oropharynx Carcinoma. *J. Clin. Oncol.* **2004**, *22*, 69–76. [CrossRef]
9. Rosenberg, A.J.; Vokes, E.E. Optimizing Treatment De-Escalation in Head and Neck Cancer: Current and Future Perspectives. *Oncologist* **2021**, *26*, 40–48. [CrossRef]
10. Ang, K.K.; Harris, J.; Wheeler, R.; Weber, R.; Rosenthal, D.I.; Nguyen-Tân, P.F.; Westra, W.H.; Chung, C.H.; Jordan, R.C.; Lu, C.; et al. Human Papillomavirus and Survival of Patients with Oropharyngeal Cancer. *N. Engl. J. Med.* **2010**, *363*, 24–35. [CrossRef]
11. Gillison, M.L.; Trotti, A.M.; Harris, J.; Eisbruch, A.; Harari, P.M.; Adelstein, D.J.; Jordan, R.C.K.; Zhao, W.; Sturgis, E.M.; Burtness, B.; et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): A randomized, multicentre, non-inferiority trial. *Lancet* **2019**, *393*, 40–50. [CrossRef] [PubMed]
12. Mehanna, H.; Robinson, M.; Hartley, A.; Kong, A.; Foran, B.; Fulton-Lieuw, T.; Dalby, M.; Mistry, P.; Sen, M.; O'Toole, L.; et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): An open-label randomized controlled phase 3 trial. *Lancet* **2019**, *393*, 51–60. [CrossRef]
13. Bigelow, E.O.; Seiwert, T.Y.; Fakhry, C. Deintensification of treatment for human papillomavirus-related oropharyngeal cancer: Current state and future directions. *Oral Oncol.* **2020**, *105*, 104652. [CrossRef]
14. Yom, S.S.; Torres-Saavedra, P.; Caudell, J.J.; Waldron, J.N.; Gillison, M.L.; Xia, P.; Truong, M.T.; Kong, C.; Jordan, R.; Le, Q.T.; et al. Faculty Opinions recommendation of Reduced-Dose Radiation Therapy for HPV-Associated Oropharyngeal Carcinoma (NRG Oncology HN002). *J. Clin. Oncol.* **2021**, *39*, 956–965. [CrossRef]
15. Siravegna, G.; O'Boyle, C.J.; Varmeh, S.; Queenan, N.; Michel, A.; Stein, J.; Thierauf, J.; Sadow, P.M.; Faquin, W.C.; Perry, S.K.; et al. Cell-Free HPV DNA Provides an Accurate and Rapid Diagnosis of HPV-Associated Head and Neck Cancer. *Clin. Cancer Res.* **2021**, *28*, 719–727. [CrossRef]
16. Ndiaye, C.; Mena, M.; Alemany, L.; Arbyn, M.; Castellsagué, X.; Laporte, L.; Bosch, F.X.; de Sanjosé, S.; Trottier, H. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: A systematic review and meta-analysis. *Lancet Oncol.* **2014**, *15*, 1319–1331. [CrossRef]
17. Neskey, D.M.; Osman, A.A.; Ow, T.J.; Katsonis, P.; McDonald, T.; Hicks, S.C.; Hsu, T.-K.; Pickering, C.R.; Ward, A.; Patel, A.; et al. Evolutionary Action Score of TP53 Identifies High-Risk Mutations Associated with Decreased Survival and Increased Distant Metastases in Head and Neck Cancer. *Cancer Res.* **2015**, *75*, 1527–1536. [CrossRef]
18. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **2015**, *517*, 576–582. [CrossRef]
19. Dassonville, O.; Formento, J.L.; Francoual, M.; Ramaioli, A.; Santini, J.; Schneider, M.; Demard, F.; Milano, G. Expression of epidermal growth factor receptor and survival in upper aerodigestive tract cancer. *J. Clin. Oncol.* **1993**, *11*, 1873–1878. [CrossRef] [PubMed]
20. Licitra, L.; Störkel, S.; Kerr, K.M.; Van Cutsem, E.; Pirker, R.; Hirsch, F.R.; Vermorken, J.B.; von Heydebreck, A.; Esser, R.; Celik, I.; et al. Predictive value of epidermal growth factor receptor expression for first-line chemotherapy plus cetuximab in patients with head and neck and colorectal cancer: Analysis of data from the EXTREME and CRYSTAL studies. *Eur. J. Cancer* **2013**, *49*, 1161–1168. [CrossRef] [PubMed]
21. Cramer, J.D.; Burtness, B.; Ferris, R.L. Immunotherapy for head and neck cancer: Recent advances and future directions. *Oral Oncol.* **2019**, *99*, 104460. [CrossRef]
22. Cottrell, T.R.; Taube, J.M. PD-L1 and Emerging Biomarkers in Immune Checkpoint Blockade Therapy. *Cancer J.* **2018**, *24*, 41–46. [CrossRef]

23. Lenouvel, D.; González-Moles, M.Á.; Talbaoui, A.; Ramos-García, P.; González-Ruiz, L.; Ruiz-Ávila, I.; Gil-Montoya, J.A. An update of knowledge on PD-L1 in head and neck cancers: Physiologic, prognostic and therapeutic perspectives. *Oral Dis.* **2020**, *26*, 511–526. [CrossRef] [PubMed]

24. Kondoh, N.; Mizuno-Kamiya, M. The Role of Immune Modulatory Cytokines in the Tumor Microenvironments of Head and Neck Squamous Cell Carcinomas. *Cancers* **2022**, *14*, 2884. [CrossRef]

25. Kim, S.Y.; Beer, M.; Vogel, D.W.T. Imaging in head and neck cancers: Update for non-radiologist. *Oral Oncol.* **2021**, *120*, 105434. [CrossRef] [PubMed]

26. Schwartz, D.L.; Harris, J.; Yao, M.; Rosenthal, D.I.; Opanowski, A.; Levering, A.; Ang, K.K.; Trotti, A.M.; Garden, A.S.; Jones, C.U.; et al. Metabolic Tumor Volume as a Prognostic Imaging-Based Biomarker for Head-and-Neck Cancer: Pilot Results from Radiation Therapy Oncology Group Protocol 0522. *Int. J. Radiat. Oncol.* **2015**, *91*, 721–729. [CrossRef] [PubMed]

27. Krupar, R.; Hautmann, M.G.; Pathak, R.R.; Varier, I.; McLaren, C.; Gaag, D.; Hellerbrand, C.; Evert, M.; Laban, S.; Idel, C.; et al. Immunometabolic Determinants of Chemoradiotherapy Response and Survival in Head and Neck Squamous Cell Carcinoma. *Am. J. Pathol.* **2018**, *188*, 72–83. [CrossRef]

28. Sandulache, V.C.; Myers, J.N. Altered metabolism in head and neck squamous cell carcinoma: An opportunity for identification of novel biomarkers and drug targets. *Head Neck* **2012**, *34*, 282–290. [CrossRef]

29. Wu, Z.-H.; Tang, Y.; Zhou, Y. A Metabolic Gene Signature to Predict Overall Survival in Head and Neck Squamous Cell Carcinoma. *Mediat. Inflamm.* **2020**, *2020*, 6716908. [CrossRef]

30. Wu, X.; Yao, Y.; Li, Z.; Ge, H.; Wang, D.; Wang, Y. Identification of a Transcriptional Prognostic Signature from Five Metabolic Pathways in Oral Squamous Cell Carcinoma. *Front. Oncol.* **2020**, *10*, 572919. [CrossRef]

31. Xing, L.; Guo, M.; Zhang, X.; Liu, F. A transcriptional metabolic gene-set based prognostic signature is associated with clinical and mutational features in head and neck squamous cell carcinoma. *J. Cancer Res. Clin. Oncol.* **2020**, *146*, 621–630. [CrossRef]

32. Carenzo, A.; Serafini, M.S.; Roca, E.; Paderno, A.; Mattavelli, D.; Romani, C.; Saintigny, P.; Koljenović, S.; Licitra, L.; De Cecco, L.; et al. Gene Expression Clustering and Selected Head and Neck Cancer Gene Signatures Highlight Risk Probability Differences in Oral Premalignant Lesions. *Cells* **2020**, *9*, 1828. [CrossRef]

33. Luo, X.-J.; Zheng, M.; Cao, M.-X.; Zhang, W.-L.; Huang, M.-C.; Dai, L.; Tang, Y.-L.; Liang, X.-H. Distinguishable Prognostic miRNA Signatures of Head and Neck Squamous Cell Cancer With or Without HPV Infection. *Front. Oncol.* **2021**, *10*, 614487. [CrossRef]

34. Bossche, V.V.D.; Zaryouh, H.; Vara-Messler, M.; Vignau, J.; Machiels, J.-P.; Wouters, A.; Schmitz, S.; Corbet, C. Microenvironment-driven intratumoral heterogeneity in head and neck cancers: Clinical challenges and opportunities for precision medicine. *Drug Resist. Updates* **2022**, *60*, 100806. [CrossRef] [PubMed]

35. Marrugo-Ramírez, J.; Mir, M.; Samitier, J. Blood-Based Cancer Biomarkers in Liquid Biopsy: A Promising Non-Invasive Alternative to Tissue Biopsy. *Int. J. Mol. Sci.* **2018**, *19*, 2877. [CrossRef] [PubMed]

36. Schmidt, H.; Kulasinghe, A.; Kenny, L.; Punyadeera, C. The development of a liquid biopsy for head and neck cancers. *Oral Oncol.* **2016**, *61*, 8–11. [CrossRef] [PubMed]

37. Pavlova, N.N.; Thompson, C.B. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab.* **2016**, *23*, 27–47. [CrossRef]

38. Xiao, C.; Eldridge, R.C.; Beitler, J.J.; Higgins, K.A.; Chico, C.E.; Felger, J.C.; Wommack, E.C.; Knobf, T.; Saba, N.F.; Shin, D.M.; et al. Association Among Glucocorticoid Receptor Sensitivity, Fatigue, and Inflammation in Patients with Head and Neck Cancer. *Psychosom. Med.* **2020**, *82*, 508–516. [CrossRef] [PubMed]

39. Accardi, C.J.P.; Walker, D.I.B.; Uppal, K.; Quyyumi, A.A.; Rohrbeck, P.D.; Pennell, K.D.; Mallon, C.T.M.; Jones, D.P. High-Resolution Metabolomics for Nutrition and Health Assessment of Armed Forces Personnel. *J. Occup. Environ. Med.* **2016**, *58* (Suppl. S8), S80–S88. [CrossRef]

40. Chandler, J.D.; Hu, X.; Ko, E.-J.; Park, S.; Lee, Y.-T.; Orr, M.; Fernandes, J.; Uppal, K.; Kang, S.-M.; Jones, D.P.; et al. Metabolic pathways of lung inflammation revealed by high-resolution metabolomics (HRM) of H1N1 influenza virus infection in mice. *Am. J. Physiol. Integr. Comp. Physiol.* **2016**, *311*, R906–R916. [CrossRef]

41. Yu, T.; Park, Y.; Li, S.; Jones, D.P. Hybrid Feature Detection and Information Accumulation Using High-Resolution LC–MS Metabolomics Data. *J. Proteome Res.* **2013**, *12*, 1419–1427. [CrossRef] [PubMed]

42. Uppal, K.; Soltow, Q.A.; Strobel, F.H.; Pittard, W.S.; Gernert, K.M.; Yu, T.; Jones, D.P. xMSanalyzer: Automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC Bioinform.* **2013**, *14*, 15. [CrossRef] [PubMed]

43. Johnson, W.; Li, C.; Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **2007**, *8*, 118–127. [CrossRef] [PubMed]

44. Liu, K.H.; Nellis, M.; Uppal, K.; Ma, C.; Tran, V.; Liang, Y.; Walker, D.I.; Jones, D.P. Reference Standardization for Quantification and Harmonization of Large-Scale Metabolomics. *Anal. Chem.* **2020**, *92*, 8836–8844. [CrossRef] [PubMed]

45. Salek, R.M.; Steinbeck, C.; Viant, M.R.; Goodacre, R.; Dunn, W.B. The role of reporting standards for metabolite annotation and identification in metabolomic studies. *Gigascience* **2013**, *2*, 1–3. [CrossRef]

46. Schymanski, E.L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H.P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48*, 2097–2098. [CrossRef]

47. Xia, J.; Wishart, D.S. MSEA: A web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res.* **2010**, *38* (Suppl. S2), W71–W77. [CrossRef]

48.   R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020. Available online: https://www.R-project.org/ (accessed on 28 February 2023).

49.   R Studio Team. *R Studio: Integrated Development Environment for R. R Studio, PBC*; R Studio Team: Boston, MA, USA, 2020. Available online: http://www.rstudio.com/ (accessed on 28 February 2023).

50.   *SEER Cancer Stat Facts: Oral Cavity and Pharynx Cancer*; National Cancer Institute: Bethesda, MD, USA, 2022. Available online: https://seer.cancer.gov/statfacts/html/oralcav.html (accessed on 28 February 2023).

51.   Zhou, J.; Xu, B.; Huang, J.; Jia, X.; Xue, J.; Shi, X.; Xiao, L.; Li, W. 1H NMR-based metabonomic and pattern recognition analysis for detection of oral squamous cell carcinoma. *Clin. Chim. Acta* **2009**, *401*, 8–13. [CrossRef]

52.   Tiziani, S.; Lopes, V.; Günther, U.L. Early Stage Diagnosis of Oral Cancer Using 1H NMR–Based Metabolomics. *Neoplasia* **2009**, *11*, 269–276, IN7–IN10. [CrossRef]

53.   Yonezawa, K.; Nishiumi, S.; Kitamoto-Matsuda, J.; Fujita, T.; Morimoto, K.; Yamashita, D.; Saito, M.; Otsuki, N.; Irino, Y.; Yoshida, M.; et al. Serum and tissue metabolomics of head and neck cancer. *Cancer Genom.-Proteom.* **2013**, *10*, 233–238.

54.   Gupta, A.; Gupta, S.; Mahdi, A.A. 1H NMR-derived serum metabolomics of leukoplakia and squamous cell carcinoma. *Clin. Chim. Acta* **2015**, *441*, 47–55. [CrossRef] [PubMed]

55.   Li, X.; Liu, L.; Li, N.; Jia, Q.; Wang, X.; Zuo, L.; Long, J.; Xue, P.; Sun, Z.; Zhao, H. Metabolomics based plasma biomarkers for diagnosis of oral squamous cell carcinoma and oral erosive lichen planus. *J. Cancer* **2022**, *13*, 76–87. [CrossRef]

56.   Li, W.; Li, Y.; Chen, Z.; Lam AK, Y.; Li, Z.; Liu, X.; Zhu, B.; Qiao, B. Analysis of metabolomics and transcriptomics data in head and neck squamous cell carcinoma. *Oral Dis.* **2022**, *29*, 1464–1479. [CrossRef] [PubMed]

57.   Jelonek, K.; Krzywon, A.; Jablonska, P.; Slominska, E.M.; Smolenski, R.T.; Polanska, J.; Rutkowski, T.; Mrochem-Kwarciak, J.; Skladowski, K.; Widlak, P. Systemic Effects of Radiotherapy and Concurrent Chemo-Radiotherapy in Head and Neck Cancer Patients—Comparison of Serum Metabolome Profiles. *Metabolites* **2020**, *10*, 60. [CrossRef] [PubMed]

58.   Rodríguez-Tomàs, E.; Arguis, M.; Arenas, M.; Royo, L.T.; Murcia, M.; Sabater, S.; Torres, L.; Baiges-Gayà, G.; Hernández-Aguilera, A.; Camps, J.; et al. Alterations in plasma concentrations of energy-balance-related metabolites in patients with lung, or head & neck, cancers: Effects of radiotherapy. *J. Proteom.* **2020**, *213*, 103605. [CrossRef]

59.   Wojakowska, A.; Zebrowska, A.; Skowronek, A.; Rutkowski, T.; Polanski, K.; Widlak, P.; Marczak, L.; Pietrowska, M. Metabolic Profiles of Whole Serum and Serum-Derived Exosomes Are Different in Head and Neck Cancer Patients Treated by Radiotherapy. *J. Pers. Med.* **2020**, *10*, 229. [CrossRef]

60.   Boguszewicz, Ł.; Bieleń, A.; Ciszek, M.; Wendykier, J.; Szczepanik, K.; Skorupa, A.; Mrochem-Kwarciak, J.; Składowski, K.; Sokół, M. NMR-Based Metabolomics in Investigation of the Radiation Induced Changes in Blood Serum of Head and Neck Cancer Patients and Its Correlation with the Tissue Volumes Exposed to the Particulate Doses. *Int. J. Mol. Sci.* **2021**, *22*, 6310. [CrossRef]

61.   Tsai, C.-K.; Lin, C.-Y.; Kang, C.-J.; Liao, C.-T.; Wang, W.-L.; Chiang, M.-H.; Yen, T.-C.; Lin, G. Nuclear Magnetic Resonance Metabolomics Biomarkers for Identifying High Risk Patients with Extranodal Extension in Oral Squamous Cell Carcinoma. *J. Clin. Med.* **2020**, *9*, 951. [CrossRef]

62.   Silva, S.; Cunha, I.; Younes, R.; Soares, F.A.; Kowalski, L.P.; Graner, E. ErbB receptors and fatty acid synthase expression in aggressive head and neck squamous cell carcinomas. *Oral Dis.* **2010**, *16*, 774–780. [CrossRef]

63.   Fhu, C.W.; Ali, A. Fatty Acid Synthase: An Emerging Target in Cancer. *Molecules* **2020**, *25*, 3935. [CrossRef]

64.   Burtness, B.; Bauman, J.E.; Galloway, T. Novel targets in HPV-negative head and neck cancer: Overcoming resistance to EGFR inhibition. *Lancet Oncol.* **2013**, *14*, e302–e309. [CrossRef]

65.   Kumar, D. Regulation of glycolysis in head and neck squamous cell carcinoma. *Postdoc J. A J. Postdr. Res. Postdr. Aff.* **2017**, *5*, 14–28. [CrossRef]

66.   Eldridge, R.C.; Uppal, K.; Hayes, D.N.; Smith, M.R.; Hu, X.; Qin, Z.S.; Beitler, J.J.; Miller, A.H.; Wommack, E.C.; Higgins, K.A.; et al. Plasma Metabolic Phenotypes of HPV-Associated versus Smoking-Associated Head and Neck Cancer and Patient Survival. *Cancer Epidemiol. Biomark. Prev.* **2021**, *30*, 1858–1866. [CrossRef]

67.   Meijer, T.W.; Kaanders, J.H.; Span, P.N.; Bussink, J. Targeting Hypoxia, HIF-1, and Tumor Glucose Metabolism to Improve Radiotherapy Efficacy. *Clin. Cancer Res.* **2012**, *18*, 5585–5594. [CrossRef]

68.   Tantiwongkosi, B.; Yu, F.; Kanard, A.; Miller, F.R. Role of18F-FDG PET/CT in pre and post treatment evaluation in head and neck carcinoma. *World J. Radiol.* **2014**, *6*, 177–191. [CrossRef] [PubMed]

69.   Gutkind, J.S.; Molinolo, A.A.; Wu, X.; Wang, Z.; Nachmanson, D.; Harismendy, O.; Alexandrov, L.B.; Wuertz, B.R.; Ondrey, F.G.; Laronde, D.; et al. Inhibition of mTOR signaling and clinical activity of metformin in oral premalignant lesions. *J. Clin. Investig.* **2021**, *6*, e147096. [CrossRef] [PubMed]

70.   Caron, A.; Richard, D.; Laplante, M. The Roles of mTOR Complexes in Lipid Metabolism. *Annu. Rev. Nutr.* **2015**, *35*, 321–348. [CrossRef] [PubMed]

71.   Chen, Y.; Qian, J.; He, Q.; Zhao, H.; Toral-Barza, L.; Shi, C.; Zhang, X.; Wu, J.; Yu, K. mTOR complex-2 stimulates acetyl-CoA and de novo lipogenesis through ATP citrate lyase in HER2/PIK3CA-hyperactive breast cancer. *Oncotarget* **2016**, *7*, 25224–25240. [CrossRef] [PubMed]

72.   Chantranupong, L.; Scaria, S.M.; Saxton, R.A.; Gygi, M.P.; Shen, K.; Wyant, G.A.; Wang, T.; Harper, J.W.; Gygi, S.P.; Sabatini, D.M. The CASTOR Proteins Are Arginine Sensors for the mTORC1 Pathway. *Cell* **2016**, *165*, 153–164. [CrossRef]

73. Sandulache, V. A Phase I/II Study of Metformin in Combination with Cisplatin and Radiation in Head and Neck Squamous Cell Carcinoma. Identifier NCT02949700. 26 January 2017–7 July 2022. Available online: https://beta.clinicaltrials.gov/study/NCT02949700 (accessed on 28 February 2023).

74. Wise-Draper, T. A Phase 2 Feasibility Study Combining Pembrolizumab and Metformin to Harness the Natural Killer Cytotoxic Response in Metastatic Head and Neck Cancer Patients. Identifier NCT04414540. 31 August 2020. Available online: https://beta.clinicaltrials.gov/study/NCT04414540 (accessed on 28 February 2023).

75. Marur, S.; Li, S.; Cmelak, A.J.; Gillison, M.L.; Zhao, W.J.; Ferris, R.L.; Westra, W.H.; Gilbert, J.; Bauman, J.E.; Wagner, L.I.; et al. E1308: Phase II Trial of Induction Chemotherapy Followed by Reduced-Dose Radiation and Weekly Cetuximab in Patients with HPV-Associated Resectable Squamous Cell Carcinoma of the Oropharynx—ECOG-ACRIN Cancer Research Group. *J. Clin. Oncol.* **2017**, *35*, 490–497. [CrossRef]

76. Couch, M.E.; Dittus, K.; Toth, M.J.; Willis, M.S.; Guttridge, D.C.; George, J.R.; Barnes, C.A.; Gourin, C.G.; Der-Torossian, H. Cancer cachexia update in head and neck cancer: Definitions and diagnostic features. *Head Neck* **2015**, *37*, 594–604. [CrossRef]

77. Boguszewicz, Ł.; Bieleń, A.; Mrochem-Kwarciak, J.; Skorupa, A.; Ciszek, M.; Heyda, A.; Wygoda, A.; Kotylak, A.; Składowski, K.; Sokół, M. NMR-based metabolomics in real-time monitoring of treatment induced toxicity and cachexia in head and neck cancer: A method for early detection of high risk patients. *Metabolomics* **2019**, *15*, 110. [CrossRef]

78. De Cecco, L.; Nicolau, M.; Giannoccaro, M.; Daidone, M.G.; Bossi, P.; Locati, L.; Licitra, L.; Canevari, S. Head and neck cancer subtypes with biological and clinical relevance: Meta-analysis of gene-expression data. *Oncotarget* **2015**, *6*, 9627–9642. [CrossRef] [PubMed]

79. Li, S.; Park, Y.; Duraisingham, S.; Strobel, F.H.; Khan, N.; Soltow, Q.A.; Jones, D.P.; Pulendran, B. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol.* **2013**, *9*, e1003123. [CrossRef] [PubMed]

80. Patel, R.M.; Roback, J.D.; Uppal, K.; Yu, T.; Jones, D.P.; Josephson, C.D. Metabolomics profile comparisons of irradiated and nonirradiated stored donor red blood cells. *Transfusion* **2015**, *55*, 544–552. [CrossRef]

81. Uppal, K.; Walker, D.I.; Jones, D.P. xMSannotator: An R Package for Network-Based Annotation of High-Resolution Metabolomics Data. *Anal Chem.* **2017**, *89*, 1063–1067. [CrossRef] [PubMed]

82. Yu, T.; Park, Y.; Johnson, J.M.; Jones, D.P. apLCMS--adaptive processing of high-resolution LC/MS data. *Bioinformatics* **2009**, *25*, 1930–1936. [CrossRef] [PubMed]