

Article

# Metabolites as Prognostic Markers for Metastatic Non-Small Cell Lung Cancer (NSCLC) Patients Treated with First-Line Platinum-Doublet Chemotherapy

Desirée Hao <sup>1,\*</sup>,<sup>†</sup> , Arjun Sengupta <sup>2,†</sup>, Keyue Ding <sup>3</sup>, ER Ubeydullah <sup>2</sup>, Saikumari Krishnaiah <sup>2</sup>, Natasha B. Leighl <sup>4</sup>, Frances A. Shepherd <sup>4</sup>, Lesley Seymour <sup>3</sup> and Aalim Weljie <sup>2,\*</sup>

<sup>1</sup> Department of Medical Oncology, Tom Baker Cancer Centre and Cummings School of Medicine, University of Calgary, 1331-29th Street N.W., Calgary, AB T2N 4N2, Canada

<sup>2</sup> Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia PA 19104, USA; arjunsen@upenn.edu (A.S.); dr.ubeydullah@gmail.com (E.U.); yksaikumari@gmail.com (S.K.)

<sup>3</sup> Canadian Cancer Trials Group, Queen's University, 10 Stuart Street, Kingston, ON K7L 3N6, Canada; kding@ctg.queensu.ca (K.D.); lseymour@ctg.queensu.ca (L.S.)

<sup>4</sup> University Health Network, Princess Margaret Cancer Centre, 700 University Avenue, Toronto, ON M5G 1Z5, Canada; natasha.leighl@uhn.ca (N.B.L.); frances.shepherd@uhn.ca (F.A.S.)

\* Correspondence: Desiree.Hao@albertahealthservices.ca (D.H.); aalim@upenn.edu (A.W.); Tel.: +1-403-521-3706 (D.H.); +1-215-573-8085 (A.W.)

† These authors contributed equally to the work.

Received: 9 June 2020; Accepted: 10 July 2020; Published: 16 July 2020



**Abstract:** The metabolic requirements of metastatic non-small cell lung (mNSCLC) tumors from patients receiving first-line platinum-doublet chemotherapy are hypothesized to imprint a blood signature suitable for survival prediction. Pre-treatment samples prospectively collected at baseline from a randomized phase III trial were assayed using nuclear magnetic resonance (NMR) spectroscopy ( $n = 341$ ) and ultra-high performance liquid chromatography – mass spectrometry (UPLC-MS) ( $n = 297$ ). Distributions of time to event outcomes were estimated by Kaplan-Meier analysis, and baseline characteristics adjusted Cox regression modeling was used to correlate markers' levels to time to event outcomes. Sixteen polar metabolites were significantly correlated with overall survival (OS) by univariate analysis ( $p < 0.025$ ). Formate, 2-hydroxybutyrate, glycine and *myo*-inositol were selected for a multivariate model. The median OS was 6.6 months in the high-risk group compared to 11.4 months in the low risk group HR (Hazard Ratio) = 1.99, 95% C.I. (Confidence Interval) 1.45–2.68;  $p < 0.0001$ ). Modeling of lipids by class (sphingolipids, acylcarnitines and lysophosphatidylcholines) revealed a median OS = 5.7 months vs. 11.9 months for the high vs. low risk group. (HR: 2.23, 95% C.I. 1.55–3.20;  $p < 0.0001$ ). These results demonstrate that metabolic profiles from pre-treatment samples may be useful to stratify clinical outcomes for mNSCLC patients receiving chemotherapy. Genomic and longitudinal measurements pre- and post-treatment may yield additional information to personalize treatment decisions further.

**Keywords:** non-small-cell-lung cancer; metabolomics; lipidomics; NMR; UPLC-MS; overall survival

## 1. Introduction

Although targeted therapies and immunotherapy have reshaped the approach to treating metastatic non-small cell lung cancer (mNSCLC), chemotherapy remains an integral part of treatment algorithms for mNSCLC. Whether chemotherapy is used alone, or in combination with novel therapies,

there is added toxicity from cytotoxic agents, thus predicting those who are most likely to benefit from the inclusion of chemotherapy remains important, even in the era of personalized medicine. Presently no prognostic or predictive biomarkers exist to identify which patients are most likely to benefit from chemotherapy.

The current study was initiated as a proof-of-principal study to evaluate metabolomics as a platform for prognostic marker discovery. To date, studies of metabolic profiling in lung cancer patients have largely focused on its potential as a diagnostic tool [1]. Various metabolites have been found at higher levels in lung cancer tissues/patients versus controls/healthy volunteers, while others appear to distinguish between tumour subtype [2] and stage [3]. In our previous pilot study evaluating serum samples collected before, during, and after chemotherapy  $\pm$  radiation, we have demonstrated that metabolite profiles varied temporally over the course of treatment [4]. Metabolite profiles have the potential to act as prognostic markers of clinical outcomes; metabolites such as hydroxylamine, tridecan-1-ol, octadecan-1-ol, were indicative of survival and metabolites such as tagatose, hydroxylamine, glucopyranose, and threonine that were reflective of progression. Building on this prior work, here we report the evaluation of metabolomic profiling in a larger, homogenous cohort of mNSCLC patients undergoing treatment with platinum-doublet chemotherapy with prospectively collected serum samples. Using a combination of nuclear magnetic resonance (NMR) and ultra-high performance liquid chromatography-mass spectrometry (UPLC-MS) lipidomics, we demonstrate that survival risk can be stratified based on the circulating levels of selected metabolites.

## 2. Results

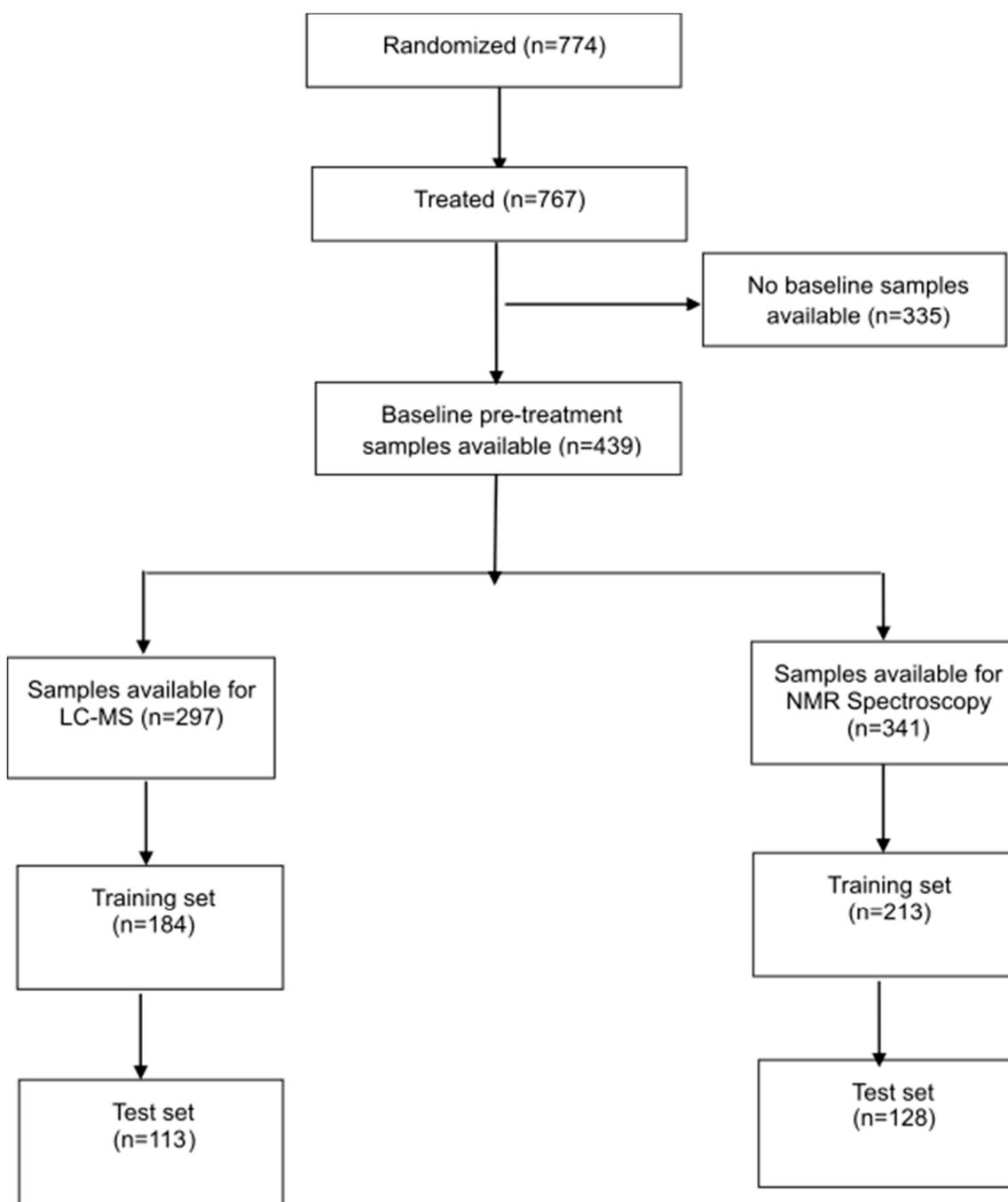
### 2.1. Patient Characteristics

Seven hundred and seventy four (774) patients from 63 participating centers in 14 countries were accrued to BR18 between April 2000 and May 2002. Full results of the clinical trial have been previously reported [5]. An interim safety analysis revealed no survival advantage for the experimental agent BMS-275291 vs. placebo groups (overall survival = 8.6 months vs. 9.2 months for BMS-275291 vs. control) leading to early termination of the study. The CONSORT diagram (Figure 1) describes the breakdown of the study population broken down into training and validation sets for NMR spectroscopy and LC-MS.

Among the 774 patients recruited, 767 patients received treatment, of which 439 had baseline pretreatment serum samples collected. 341 and 297 serum samples were used for NMR spectroscopy and LC-MS respectively. Patients' clinical characteristics by analytic method (NMR vs. LC-MS) are outlined in Table 1; Table 2. There were no significant differences in age, gender, performance status, treatment received or OS between patients with serum samples available for analysis and the overall study population.

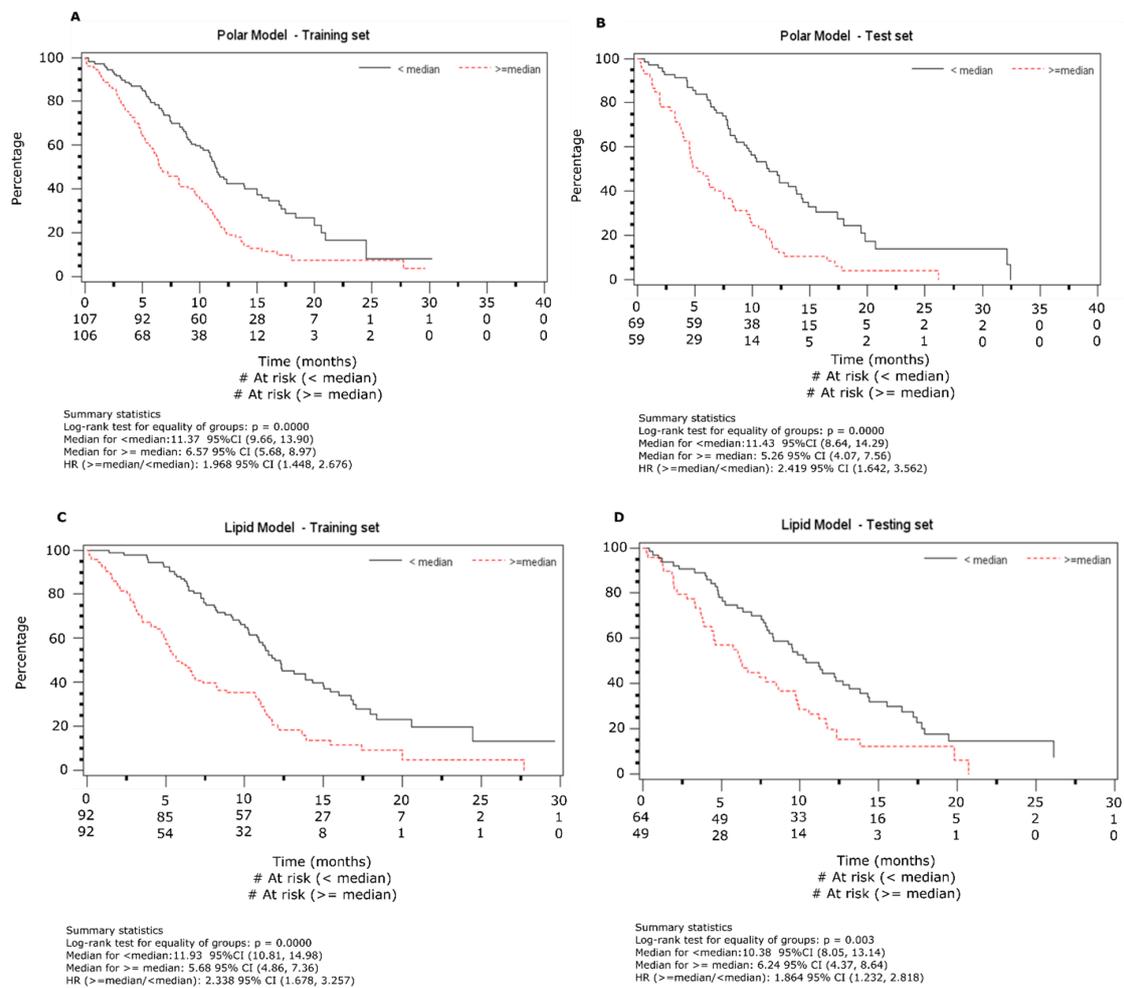
### 2.2. Overall Survival Analysis from NMR Data

NMR spectroscopy was used to quantitatively profile water soluble polar molecules of low molecular weight. Among the 341 samples in which NMR data was acquired, 213 were used as a training set and 128 in the test set. Concentration data of each metabolite was  $\log_2$  transformed and z-scored by mean centering. Out of 43 polar metabolites, 16 were found to be significantly correlated with the overall survival rate using univariate cox regression (Table S1,  $p < 0.025$ ). Final multivariate prognostic model was created from 4 metabolites (2-hydroxybutyrate, formate, glycine and *myo*-inositol) after backward variable elimination at 5% significance level (Table S1).



**Figure 1.** Study consort diagram describing the study population breakdown into training and validations sets for NMR spectroscopy and LC-MS.

Based on the model, each patient was assigned a prognostic score and were divided into two groups based on the median prognostic score. Median overall survival (OS) was found to be 11.4 months for the low risk and 6.6 months for the high-risk group (HR = 1.99, 95% C.I 1.45–2.68;  $p < 0.0001$ , Figure 2A). Multivariable Cox regression model adjusted for baseline factors was performed to check the high-risk group's effect and found its effect is similar to the univariate analysis (HR: 2.16, 95% C.I. 1.57–2.98;  $p < 0.0001$ , Table 1).



**Figure 2.** Prognostic survival analysis using the nuclear magnetic resonance (NMR) and lipid data sets. The data is presented using Kaplan-Meier curves on test and training sets from the two analytical platforms. Samples were divided into training and test sets as described in Figure 1. (A). NMR training dataset, (B). NMR test dataset, (C). Lipid training dataset and (D). Lipid test dataset.

**Table 1.** Prognostic effects of polar metabolite markers on overall survival after adjusting for baseline factors.

A. Polar Metabolites Training Set			
Factors	Hazard ratio	95% C.I. of HR	p-Value
Risk group			
High	2.16	1.57–2.97	<0.0001
Low	1		
ECOG Performance status			
2, 3	1.47	0.89–2.41	0.13
0, 1	1		
Weight loss			
≥5%	1.89	1.35–2.66	0.0004
<5%	1		

Table 1. Cont.

A. Polar Metabolites Training Set			
Factors	Hazard ratio	95% C.I. of HR	<i>p</i> -Value
Sex			
Male	1.30	0.87–1.94	0.20
Female	1		
Disease Stage			
III	1.20	0.81–1.78	0.35
IV	1		
Histology			
Squamous	0.99	0.70–1.42	0.98
Non-Squamous	1		
Hemoglobin <sup>1</sup>			
Grade 1+	1.51	1.18–2.23	0.01
Grade 0	1		
B. Polar Metabolites Test Set			
Factors	Hazard ratio	95% C.I. of HR	<i>p</i> -Value
Risk group			
High	2.42	1.61–3.65	<0.0001
Low	1		
ECOG Performance status			
2, 3	1.74	0.89–3.42	0.11
0, 1	1		
Weight loss			
≥5%	1.48	0.93–2.34	0.10
<5%	1		
Sex			
Male	0.98	0.61–1.59	0.94
Female	1		
Disease Stage			
III	0.51	0.30–0.85	0.01
IV	1		
Histology			
Squamous	0.97	0.59–1.59	0.90
Non-Squamous	1		
Hemoglobin <sup>1</sup>			
Grade 1+	1.59	1.06–2.68	0.04
Grade 0	1		

<sup>1</sup>: Within normal limits (grade 0), <Lower limit of normal (grade 1+). H.R.: Hazard ratio, C.I.: confidence interval. ECOG Eastern Cooperative Oncology Group.

The test set was used for validating the results found from the training set. In order to do so, the risk scores for patients in the test set were separated into two groups based on the median of the training set. The Kaplan-Meier (K-M) estimate of the survival distribution based on the median is presented in Figure 2B. The median OS are 11.6 and 5.3 months for the low and high-risk groups, respectively (HR = 2.43, 95% C.I. 1.61–3.66;  $p < 0.0001$ ). Multivariate Cox regression modeling, adjusted for baseline factors was performed to check the effect of the high risk group, and found to be similar to the univariate analysis (HR: 2.43, 95% C.I. 1.61–3.66;  $p < 0.0001$ , Table 1).

Among the four metabolites used for constructing the multivariable model, 2-hydroxybutyrate (2-HB,  $p = 0.001$ ) and formate ( $p = 0.002$ ) were found to be significantly associated to the OS in the

test set. The correlation between two other metabolites, glycine ( $p = 0.11$ ) and *myo*-inositol ( $p = 0.31$ ), trended the same direction, but did not meet the significance threshold.

### 2.3. Overall Survival Analysis from Lipid Data

A total of 297 baseline lipid profiles were assayed. The dataset was randomly divided into two subsets with roughly 3:2 ratio (184 patients in training set and 113 in testing set). The data was  $\log_2$  transformed and mean centered to construct z scores. A total of 1420 lipid features were analyzed via univariate Cox regression modeling to establish each variable's correlation with OS from training set data and 53 variables (Table S2,  $p < 0.025$ ) were identified. To take care of the inter-variable co-linearity due to biosynthetic pathways and other factors, these variables were subjected to hierarchical cluster analysis resulting in 12 branches (Table S3). For the branches with moderate to significantly correlated lipid features, principal component analysis (PCA) was carried out and the first principal component (PC) from each such branch was used as the representative variable for further analysis, 9 PCs were identified using this method. Backward variable elimination with a 5% significance level led final to a multivariable model using 4 variables — two of those include largely sphingolipids and triglycerides while specific lipid classes could not be assigned to the other two classes (Table S3). A prognostic score for each patient was generated as detailed in Supplementary Information S1. Based on these scores the patients were separated into high and low risk groups. Using Cox regression modeling, the median OS for the high-risk group was found to be 5.7 months while median OS for low risk group was 11.9 months (Figure 2C, HR: 2.34, 95% C.I. 1.68–3.26;  $p < 0.0001$ ). Multivariate Cox regression modeling adjusted for baseline factors was performed and results were similar to the univariate analysis (Table 2, HR: 2.23, 95% C.I. 1.55–3.20;  $p < 0.0001$ ).

**Table 2.** Prognostic effects of lipid markers on overall survival after adjusting for baseline factors.

A. Lipids Training Set			
Factors	Hazard ratio	95% C.I. of HR	<i>p</i> -Value
Risk group			
High	2.23	1.55–3.22	<0.0001
Low	1		
ECOG Performance status			
2, 3	1.51	0.87–2.64	0.14
0, 1	1		
Weight loss			
≥5%	1.93	1.31–2.85	0.0009
<5%	1		
Sex			
Male	1.62	1.02–2.56	0.04
Female	1		
Disease Stage			
III	1.20	0.80–1.81	0.38
IV	1		
Histology			
Squamous	0.95	0.60–1.49	0.81
Non-Squamous	1	1	
Hemoglobin <sup>1</sup>			
Grade 1+	1.52	1.12–2.34	0.03
Grade 0	1		

Table 2. Cont.

B. Lipids Test Set			
Factors	Hazard ratio	95% C.I. of HR	p-Value
Risk group			
High	1.83	1.19–3.22	0.006
Low	1		
ECOG Performance status			
2, 3	1.70	0.75–3.86	0.21
0, 1	1		
Weight loss			
≥5%	1.55	0.86–2.79	0.15
<5%	1		
Sex			
Male	1.11	0.61–2.02	0.83
Female	1		
Disease Stage			
III	0.58	0.32–1.08	0.09
IV	1		
Histology			
Squamous	0.85	0.45–1.61	0.62
Non-Squamous	1		
Hemoglobin <sup>1</sup>			
Grade 1+	1.60	1.03–2.68	0.04
Grade 0	1		

<sup>1</sup>: Within normal limits (grade 0), <Lower limit of normal (grade 1+). H.R.: Hazard Ratio; C.I: Confidence Interval. ECOG Eastern Cooperative Oncology Group.

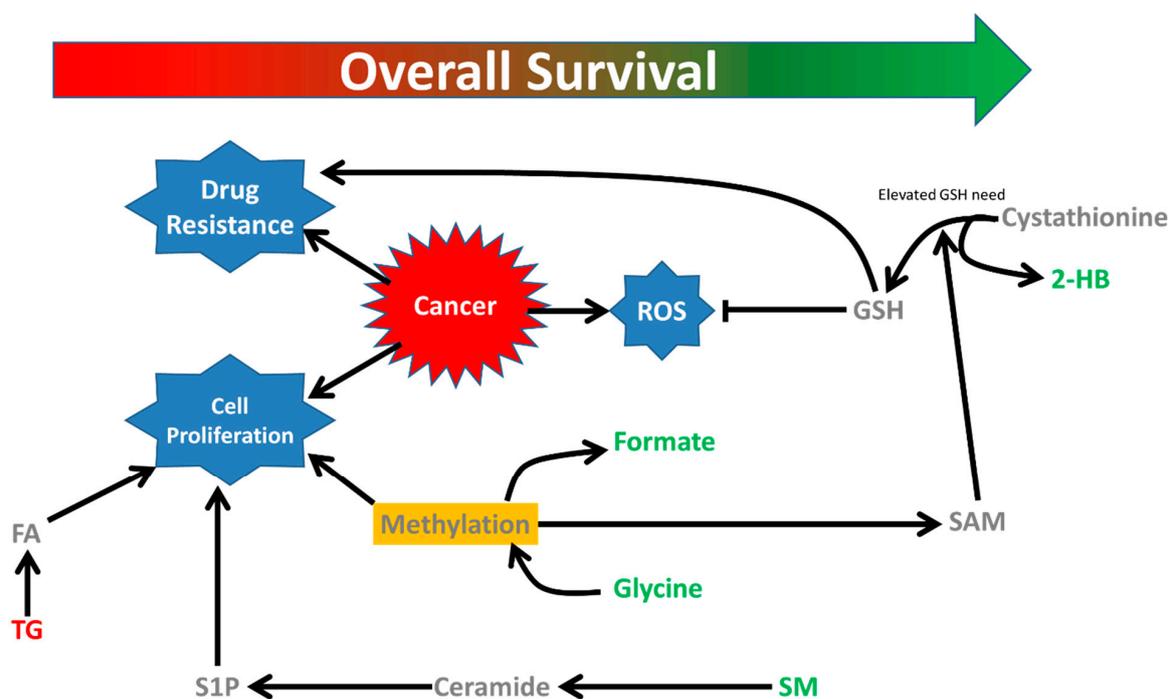
In the lipid data generated from LC-MS, four PC variables were used to create a risk score and patients were divided into high and low risk groups based on the median risk scores. The K-M estimate of the OS was found to be 10.4 and 6.2 months for the low and high-risk groups, respectively (Figure 2D, HR: 1.86, 95% C.I. 1.23–2.82;  $p = 0.003$ ). As with the polar data, multivariable Cox regression modeling adjusted for baseline factors was performed to evaluate the effect of the high risk group; results were similar to the univariate analysis (Table 2, HR: 1.83, 95% C.I. 1.19–2.81;  $p = 0.006$ ).

### 3. Discussion

Pre-treatment serum metabolites and lipid profiles in mNSCLC patients have the potential to serve as prognostic markers of clinical outcomes. Using prospectively collected serum samples from a large cohort of mNSCLC that was linked to outcomes, we demonstrated that with both polar data and lipid data, patients could be divided into high and low risk groups with differential OS. The results suggest that NMR and LC-MS characterize different, but complementary, metabolomic and lipidomic profiles that each identifies a high vs. low risk group even after correcting for known clinical prognostic variable such as gender, stage, histology and performance status. Tian et al. [6] recently reported analogous findings using pre-treatment serum metabolic profiles from a larger cohort of mNSCLC patients treated with platinum-doublet chemotherapy, providing support for this approach. The results from our study are complementary. The metabolite panel reported by Tian et al. was associated with longer median progression free survival, but was a different set of seven metabolites (hypotaurine, uridine, dodecanoylcarnitine, choline, dimethylglycine, niacinamide, L-palmitoylcarnitine) [6] perhaps reflecting differences in the patient population or the systemic therapy used.

Results from our polar NMR data suggest that elevation of glutathione synthesis supported by elevated methylation pathways may be associated with better survival in NSCLC patients (Figure 3).

We observed that elevated blood 2-HB, glycine and formate, were all positively associated with overall survival (Table S1). 2-HB is a metabolite related to glutathione synthesis [7]. Glutathione regulates the activity of ribonucleotide reductase that synthesizes deoxyribonucleotides from ribonucleotides [8] which is required for both cancer initiation and proliferation [9]. In NSCLC patients, elevated glutathione-S-transferase (GST) expression has been shown to decrease chemotherapeutic response [10,11]. The balance between the opposing effects of glutathione synthesis and metabolism may regulate the outcome of the disease. Glycine is also key component of glutathione synthesis [12].



**Figure 3.** Metabolic changes associated with overall survival in NSCLC patients. Alteration in glutathione metabolic pathway is apparent from association of 2-hydroxybutyrate and survival. In addition, methylation pathways (formate, glycine), triglyceride and sphingomyelin metabolism are also associated to overall survival. FA: fatty acid, TG: triacylglycerol, S1P: Sphingosine-1-phosphate; SM: Sphingomyelin, SAM: S-adenosylmethionine, GSH: reduced glutathione, ROS: Reactive oxygen species, 2-HB: 2-hydroxybutyrate.

Both glycine and formate are byproducts of transmethylation via serine which donates the one carbon unit to tetrahydrofolate that further leads to purine/pyrimidine synthesis and transmethylation via the methionine cycle [13]. Two of the metabolites from Tian et al. are also involved in one-carbon metabolism (dimethylglycine, choline) [6]. Of note, another study designed to elucidate metabolites predictive for platinum-chemotherapy efficacy primarily identified caffeine-based metabolites [14] which affects plasma homocysteine level [15], a key reporter metabolite for one carbon metabolism. Hypermethylation facilitates cancer cell proliferation in various ways including reduced tumor suppressor gene expression [16], regulation of translation by methylation of RNA [17] and regulation of protein function by posttranslational modification [13].

Recent studies have shown that serine and glycine can provide the adenosine moiety needed for synthesis of S-adenosylmethionine (SAM) from methionine to support the methylation cycle [18] and in turn generate homocysteine which can lead to excessive systemic glutathione. The positive association of overall survival and methylation metabolites, glycine and formate, may ultimately be linked to glutathione metabolism.

The lipid analysis of the patient serum samples provides further interesting insights to survival in NSCLC. Sphingolipids were positively associated to overall survival (Table S2). One of the three sphingolipids that was positively associated to overall survival was sphingomyelin (SM d18:2/25:0). The other two sphingolipids were putatively identified as ceramides based on their retention times. We hypothesize that elevated ceramide synthesis from membrane sphingomyelin may be associated to less aggressive tumorigenesis and better overall survival. Ceramide, a central sphingolipid which may be synthesized from sphingomyelin by sphingomyelinase [19], is associated with cell growth inhibition, induction of apoptosis, autophagy and ER stress response [20] and thus may be considered a tumor repressor lipid. We also observed several triglycerides were negatively associated to the overall survival. Cancer cells growth is enhanced in lipid-rich environments where the triglycerides are hydrolyzed to provide the fatty acids that may be used as the energy source by the cells [21]. Limiting fatty acids availability could serve as a potential therapeutic strategy against cancer [22], thus the negative association of triglyceride level with overall survival seen in this study reflects may reflect limited availability of fatty acids essential for the growth of cancer cells ultimately leading to improved survival of the patients.

Other metabolomics studies in cancer patients have used different analytic platforms such as gas chromatography-mass spectrometry (GC-MS) and LC-MS. <sup>1</sup>H-NMR can measure a wide range of metabolites with little sample preparation but is limited by a lower sensitivity and requires more expensive instrumentation. However, the technique is highly quantitative and hence the diagnostic value of biomarkers identified is much more useful from a clinical perspective. In contrast, hyphenated mass spectrometry generally requires more extensive sample preparation and can usually measure only a specific subset of metabolites depending on the type of hyphenation techniques that precedes the detection (GC/CE/LC). For example, hydrophilic interaction chromatography (HILIC) is more suitable for polar metabolites while reverse phase chromatography detects lipids/nonpolar metabolites more efficiently. The importance of polar metabolites in progression and prediction of survival in lung cancer have been underscored by different studies [14,23,24]. However, despite significant evidence of alteration in lipid metabolism in lung cancer [25], untargeted lipidomics analyses, as was performed in our study, have rarely been employed for biomarker discovery in lung cancer patients. NMR spectroscopy is not a suitable technique for measuring lipids and mass spectrometry coupled with reverse phase chromatography offers a more suitable approach.

The main strength of our study was examined a large, homogenous population of patients with mNSCLC who were treated with first line cytotoxic chemotherapy. The serum samples we analyzed were uniformly collected prior to receiving systemic therapy hence the treatment would not have influenced our findings. Although patients in Canadian Clinical Trials Group (CCTG) BR18 study were treated with BMS-275291 vs placebo, the addition of BMS-275291 did not affect survival [5]; therefore our results serve as proof-of-principal that in mNSCLC patients receiving platinum-doublet chemotherapy, the pre-treatment metabolite profile could potentially be used as a prognostic marker. However, because the serum samples we evaluated were collected from a cohort dating back to 2000 to 2002, at a time when genomic profiling and/or tumour PDL-1 (Programmed death-ligand 1) status was not evaluated, it is conceivable that imbalances in tumour characteristics might have accounted for the observed differences between high and low risk groups. Confirmation in a more contemporary cohort would be necessary to substantiate these findings.

## 4. Methods and Materials

### 4.1. Patients

All serum samples used were prospectively collected from patients with mNSCLC participating in the CCTG BR 18 study, a randomized, double-blind phase III trial evaluating carboplatin, paclitaxel plus either placebo vs. the matrix metalloproteinase inhibitor BMS-275291 [5]. The samples were all collected and processed in a uniform fashion as per protocol. All samples were stored at the CCTG

Tumour Tissue Data Repository in Kingston (Kingston ON, Canada). Clinical data including age, gender, stage, histology, and outcome were collected as part of the trial. This translational study was approved by the Health Research Ethics Committee of Alberta-Cancer Committee and the University of Pennsylvania.

As part of the original BR-18 trial, patients provided written informed consent with regards to participation in the trial and collection of blood samples for future studies.

This study was approved by the University of Calgary Conjoint Health Research Ethics Board (Ethics ID: E-24705, 2012-08-22) and the University of Pennsylvania Institutional Review Board (Protocol #816578, 2012-09-27).

#### 4.2. Sample Preparation

We used two different methodologies to evaluate metabolomics signatures: nuclear magnetic resonance ( $^1\text{H-NMR}$ ) spectroscopy and liquid chromatography mass spectrometry (LC-MS). Blood serum (200  $\mu\text{L}$ ) was extracted using 1:2 chloroform/methanol. The upper polar fraction was dried using vacuum concentrator and the lower non-polar fraction was dried overnight in the hood. The polar fraction was further dissolved in 200  $\mu\text{L}$  NMR solvent made up in 10%  $\text{D}_2\text{O}$  with pH adjusted to  $\sim 7.0$  by  $\text{NaHPO}_4/\text{NaH}_2\text{PO}_4$  and 0.25 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) added as internal standard. The reconstituted samples were put into 3 mm NMR tubes (Bruker Biospin, Billerica, MA, USA) in 96 tube racks designed for samplejet (Bruker Biospin) for acquiring the NMR spectra.

The non-polar fraction was used for analyzing the lipids. 40  $\mu\text{L}$  of the fraction was dried and dissolved in 700  $\mu\text{L}$  of solvent containing 60% solvent A (40%  $\text{H}_2\text{O}$ , 60% ACN, 10 mM ammonium formate) and 40% solvent B (90% isopropanol 10% acetonitrile, 10 mM ammonium formate). 30  $\mu\text{L}$  of the reconstituted solvent was added to microtubes containing 10  $\mu\text{L}$  of internal standards (detail needed). The sample was further diluted to 150  $\mu\text{L}$  using 60% A/40% B for injection.

#### 4.3. NMR Spectroscopy

All NMR spectra were acquired in a 700 MHz Bruker AVANCE III HD NMR spectrometer equipped with a 3 mm triple resonance inverse probe and Samplejet automated sample handling system (Bruker Biospin, Billerica, MA, USA). The pulseprogram used for acquiring the spectra took the shape of first transient of 2-dimensional NOESY spectroscopy (RD-90-t-90- $t_m$ -90-ACQ, where RD = relaxation delay, t = echo time,  $t_m$  = mixing time and ACQ = acquisition) [26]. Spectral acquisition was performed in Topspin 3.0 (Bruker Biospin). The FIDs (free induction decay) were acquired using 96 scans each acquiring 76K data points from a spectral width of 14 ppm with acquisition time 4 s for each FID. The FIDs generated was further exported for targeted NMR spectral profiling:

Targeted spectral profiling [27] was performed using Chenomx v 8.0 (Chenomx Inc. Edmonton, AB, Canada). Imported FIDs were zero filled to 128k and applied line broadening factor 0.1 Hz and Fourier transformed. All spectra were calibrated with respect to the internal standard (0.25 mM DSS) and the spectral peaks were fitted with Chenomx reference library.

#### 4.4. Lipidomics Assay by UPLC-Qtof-MS

Dried lipid fraction was reconstituted in 60% solvent A (40%  $\text{H}_2\text{O}$ , 60% ACN, 10 mM ammonium formate) and 40% solvent B (90% IPA, 10% ACN, 10 mM ammonium formate). These samples were transferred into glass vials for ultra-performance liquid chromatography coupled with a qTOF Xevo G2S detector (Waters Corporation, Milford, MA, USA) for high throughput LC-MS-based lipidomics. 10  $\mu\text{L}$  sample was injected into a reverse phase column (XSELECTTM CSHTM C18, 2.1 mm  $\times$  100 mm  $\times$  2.5  $\mu\text{m}$ ) using an Aquity H-class UPLC system (Waters Corporation). Samples were chromatographed for 9 min at 0.5 mL/min flow rate. The UPLC gradient was as follows - 75% A and 25% B for 0.5 min, a quick ramp of 50% A and 50% B for 0.5 min, 25% A and 75% B for 4 min, followed by a ramp of 10% A and 90% B for 2 min, and finally a ramp to 1% A and 99% B for 2 min.

Column eluent was directly introduced into mass spectrometer that was performed in both positive and negative ion-sensitive mode with a capillary voltage of 3000 V and a sampling cone voltage of 40 °C. The desolvation gas flow was set to 800 L/h, and the temperature was set to 450 °C. The source temperature was set to 80 °C. Assessment of accurate mass was maintained by the introduction of a lock-spray interface of leucine-enkephalin (556.2771 m/z) at a concentration of 0.5 ng/μL in 50% aqueous acetonitrile and a rate of 5 μL/min. Data was acquired in the centroid MSe mode from 50–1200 m/z mass ranges for both MS (low energy) and MSe (high energy) modes. Low-energy or fragmented data were collected without collision energy, whereas high-energy or fragmented data were collected by using a collision-energy ramp from 15–40 eV. The entire set of duplicate sample injection was bracketed with test mix of standard metabolites at the beginning and end of run for evaluating instrument performance. Samples were randomized and injected in duplicates with pooled quality control sample injection after every 10 runs to compensate for instrumental drift. Data analysis and lipid identification were performed with Progenesis QI 2.3 software (Waters Corporation, Milford, MA, USA).

#### 4.5. Statistical Analysis

Exploratory analyses were performed to characterize the relationships between patients' lipid and polar metabolite levels with baseline characteristics and outcomes. Chi-square test was used to assess association between categorical variables; Principle component analysis was used to synthesize information of correlated variables. Kaplan-Meier curves were used to estimate the distributions of time to event outcomes, log-rank test was used to test difference between groups, and Cox regression model was used to correlate markers' levels to time to event outcomes while adjusting baseline characteristics.

Prior to any analysis, patients with NMR and/or LC-MS data were randomly divided in approximately a 3:2 ratio into training and testing sets, stratified by treatment received (BMS275291 vs. Placebo), stage (III vs. IV), and ECOG PS (0, 1, vs. ≥2). Overall survival between the two cohorts were compared and found to be similar. We then used the 3/5 of the data as training set, and the remaining 2/5 of the data as test set.

## 5. Conclusions

In an era of rapidly changing standards of care that now include immunotherapy and targeted therapy, conventional cytotoxic chemotherapy still forms part of the treatment algorithm for most mNSCLC patients. Therefore, tools to guide clinicians regarding which patients are the most likely to benefit from the addition of chemotherapy remain relevant in current times. Our results align with other metabolomics studies and taken together, suggest that this minimally invasive strategy warrants further validation among mNSCLC. In addition, the evaluation of changes in metabolomics/lipidomic profiles over the course of treatment may provide further insights into which patients are most likely to benefit for a specific therapeutic approach.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2072-6694/12/7/1926/s1>, Table S1: Polar metabolites that significantly correlated to overall survival in the training set measured by NMR, Table S2: Lipids that significantly correlated to overall survival in the training set measured by RP-UPLC-qTOF-MS, Table S3: 53 significant lipid features were subjected to hierarchical cluster analysis to compensate for the linearity due to chemical similarity, Supplementary information S1: Statistical methodologies for analyzing correlated lipid data.

**Author Contributions:** Conceptualization, D.H., A.W., F.A.S., L.S., and N.B.L.; Methodology, A.S., E.U., S.K., A.W.; Formal Analysis, K.D., Resources, A.W.; Writing—Original Draft Preparation, A.S., D.H., A.W.; Writing—Review & Editing, A.S., D.H., A.W.; Visualization, A.S., K.D., D.H.; Supervision, D.H., A.W.; Funding Acquisition, A.W., D.H., K.D., L.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Alberta Cancer Foundation and Alberta Innovates Health Solutions grants and NIH grant R21-CA-213234 to AMW; the Canadian Cancer Society Grant # 704970 to K.D. and L.S.

**Conflicts of Interest:** Authors declare no conflict of interest.

**Data Availability:** The data is available at [www.doi.org/10.5281/zenodo.3887012](http://www.doi.org/10.5281/zenodo.3887012).

## References

1. Hori, S.; Nishiumi, S.; Kobayashi, K.; Shinohara, M.; Hatakeyama, Y.; Kotani, Y.; Hatano, N.; Maniwa, Y.; Nishio, W.; Bamba, T.; et al. A metabolomic approach to lung cancer. *Lung Cancer* **2011**, *74*, 284–292. [[CrossRef](#)]
2. Jordan, K.W.; Adkins, C.B.; Su, L.; Halpern, E.F.; Mark, E.J.; Christiani, D.C.; Cheng, L.L. Comparison of squamous cell carcinoma and adenocarcinoma of the lung by metabolomic analysis of tissue-serum pairs. *Lung Cancer* **2010**, *68*, 44–50. [[CrossRef](#)]
3. Deja, S.; Porebska, I.; Kowal, A.; Zabek, A.; Barg, W.; Pawelczyk, K.; Stanimirova, I.; Daszykowski, M.; Korzeniewska, A.; Jankowska, R.; et al. Metabolomics provide new insights on lung cancer staging and discrimination from chronic obstructive pulmonary disease. *J. Pharm. Biomed. Anal.* **2014**, *100*, 369–380. [[CrossRef](#)] [[PubMed](#)]
4. Hao, D.; Sarfaraz, M.O.; Farshidfar, F.; Bebb, D.G.; Lee, C.Y.; Card, C.M.; David, M.; Weljie, A.M. Temporal characterization of serum metabolite signatures in lung cancer patients undergoing treatment. *Metabolomics* **2016**, *12*, 1–9. [[CrossRef](#)] [[PubMed](#)]
5. Leighl, N.B.; Paz-Ares, L.; Douillard, J.-Y.; Peschel, C.; Arnold, A.; Depeirre, A.; Santoro, A.; Betticher, D.C.; Gatzemeier, U.; Jassem, J.; et al. Randomized Phase III Study of Matrix Metalloproteinase Inhibitor BMS-275291 in Combination With Paclitaxel and Carboplatin in Advanced Non-Small-Cell Lung Cancer: National Cancer Institute of Canada-Clinical Trials Group Study BR.18. *J. Clin. Oncol.* **2005**, *23*, 2831–2839. [[CrossRef](#)] [[PubMed](#)]
6. Tian, Y.; Wang, Z.; Liu, X.; Duan, J.; Feng, G.; Yin, Y.; Gu, J.; Chen, Z.; Gao, S.; Bai, H.; et al. Prediction of chemotherapeutic efficacy in non-small cell lung cancer by serum metabolomic profiling. *Clin. Cancer Res.* **2018**, *24*, 2100–2109. [[CrossRef](#)]
7. Gall, W.E.; Beebe, K.; Lawton, K.A.; Adam, K.P.; Mitchell, M.W.; Nakhle, P.J.; Ryals, J.A.; Milburn, M.V.; Nannipieri, M.; Camastra, S.; et al.  $\alpha$ -hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. *PLoS ONE* **2010**, *5*, e10883. [[CrossRef](#)]
8. Sengupta, R. Thioredoxin and glutaredoxin-mediated redox regulation of ribonucleotide reductase. *World J. Biol. Chem.* **2014**, *5*, 68–74. [[CrossRef](#)] [[PubMed](#)]
9. Harris, I.S.; Treloar, A.E.; Inoue, S.; Sasaki, M.; Gorrini, C.; Lee, K.C.; Yung, K.Y.; Brenner, D.; Knobbe-Thomsen, C.B.; Cox, M.A.; et al. Glutathione and Thioredoxin Antioxidant Pathways Synergize to Drive Cancer Initiation and Progression. *Cancer Cell* **2015**, *27*, 211–222. [[CrossRef](#)] [[PubMed](#)]
10. Inoue, T.; Ishida, T.; Sugio, K.; Maehara, Y.; Sugimachi, K. Glutathione S transferase Pi Is a powerful indicator in chemotherapy of human lung squamous-cell carcinoma. *Respiration* **1995**, *62*, 223–227. [[CrossRef](#)]
11. Bai, F.; Nakanishi, Y.; Kawasaki, M.; Takayama, K.; Yatsunami, J.; Pei, X.H.; Tsuruta, N.; Wakamatsu, K.; Hara, N. Immunohistochemical expression of glutathione S-transferase-Pi can predict chemotherapy response in patients with nonsmall cell lung carcinoma. *Cancer* **1996**, *78*, 416–421. [[CrossRef](#)]
12. Lu, S.C. Glutathione Synthesis. *Biochim. Biophys. Acta* **2014**, *1830*, 3143–3153. [[CrossRef](#)] [[PubMed](#)]
13. Newman, A.; Maddocks, O. One-carbon metabolism in cancer. *Br. J. Cancer* **2017**, *116*, 1499–1504. [[CrossRef](#)] [[PubMed](#)]
14. Shen, J.; Ye, Y.; Chang, D.W.; Huang, M.; Heymach, J.V.; Roth, J.A.; Wu, X.; Zhao, H. Circulating metabolite profiles to predict overall survival in advanced non-small cell lung cancer patients receiving first-line chemotherapy. *Lung Cancer* **2017**, *114*, 70–78. [[CrossRef](#)] [[PubMed](#)]
15. Verhoef, P.; Pasman, W.J.; Van Vliet, T.; Urgert, R.; Katan, M.B. Contribution of caffeine to the homocysteine-raising effect of coffee: A randomized controlled trial in humans. *Am. J. Clin. Nutr.* **2002**, *6*, 1244–1248. [[CrossRef](#)] [[PubMed](#)]
16. Kulis, M.; Esteller, M. 2-DNA Methylation and Cancer. *Adv. Genet.* **2010**, *70*, 27–56. [[CrossRef](#)]
17. Fu, Y.; Dominissini, D.; Rechavi, G.; He, C. Gene expression regulation mediated through reversible m<sup>6</sup>A RNA methylation. *Nat. Rev. Genet.* **2014**, *15*, 293–306. [[CrossRef](#)]
18. Maddocks, O.; Labuschagne, C.; Adams, P.; Vousden, K. Serine metabolism supports the methionine cycle and DNA/RNA Methylation through de novo ATP synthesis in cancer cells. *Mol. Cell* **2016**, *61*, 210–221. [[CrossRef](#)]
19. Mullen, T.D.; Hannun, Y.A.; Obeid, L.M. Ceramide synthases at the centre of sphingolipid metabolism and biology. *Biochem. J.* **2012**, *441*, 789–802. [[CrossRef](#)]

20. Ponnusamy, S.; Meyers-Needham, M.; Senkal, C.E.; Saddoughi, S.A.; Sentelle, D.; Selvam, S.P.; Salas, A.; Ogretmen, B. Sphingolipids and cancer: Ceramide and sphingosine-1-phosphate in the regulation of cell death and drug resistance. *Future Oncol.* **2010**, *6*, 1603–1624. [[CrossRef](#)]
21. Beloribi-Djefaflija, S.; Vasseur, S.; Guillaumond, F. Lipid metabolic reprogramming in cancer cells. *Oncogenesis* **2016**, *5*, e189. [[CrossRef](#)] [[PubMed](#)]
22. Currie, E.; Schulze, A.; Zechner, R.; Walther, T.C.; Farese, R.V. Cellular fatty acid metabolism and cancer. *Cell Metab.* **2013**, *18*, 153–161. [[CrossRef](#)] [[PubMed](#)]
23. Puchades-Carrasco, L.; Jantus-Lewintre, E.; Pérez-Rambla, C.; García-García, F.; Lucas, R.; Calabuig, S.; Blasco, A.; Dopazo, J.; Camps, C.; Pineda-Lucena, A. Serum metabolomic profiling facilitates the non-invasive identification of metabolic biomarkers associated with the onset and progression of non-small cell lung cancer. *Oncotarget* **2016**, *7*, 12904–12916. [[CrossRef](#)] [[PubMed](#)]
24. Musharraf, S.G.; Mazhar, S.; Choudhary, M.I.; Rizi, N. Plasma metabolite profiling and chemometric analyses of lung cancer along with three controls through gas chromatography-mass spectrometry. *Sci. Rep.* **2015**, *5*, 8607. [[CrossRef](#)]
25. Salvador, M.M.; de Cedrón, M.G.; Rubio, J.M.; Martínez, S.F.; Martínez, R.S.; Casado, E.; de Molina, A.R.; Sereno, M. Lipid metabolism and lung cancer. *Crit. Rev. Oncol. Hematol.* **2017**, *112*, 31–40. [[CrossRef](#)]
26. Beckonert, O.; Keun, H.C.; Ebbels, T.M.D.; Bundy, J.; Holmes, E.; Lindon, J.C.; Nicholson, J.K. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Prot.* **2007**, *2*, 2692–2703. [[CrossRef](#)]
27. Weljie, A.M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C.M. Targeted Profiling: Quantitative Analysis of <sup>1</sup>H NMR Metabolomics Data. *Anal. Chem.* **2006**, *78*, 4430–4442. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).