Review

Plasma Metabolite Profiling in the Search for Early-Stage Biomarkers for Lung Cancer: Some Important Breakthroughs

Jill Meynen [1], Peter Adriaensens [2,*], Maarten Criel [1,3], Evelyne Louis [4], Karolien Vanhove [1,4,5], Michiel Thomeer [3], Liesbet Mesotten [1,6] and Elien Derveaux [2]

1   Faculty of Medicine and Life Sciences, Hasselt University, Martelarenlaan 42, B-3500 Hasselt, Belgium; jill.meynen@uhasselt.be (J.M.); maarten.criel@zol.be (M.C.); karolien.vanhove@uzleuven.be (K.V.); liesbet.mesotten@zol.be (L.M.)
2   Applied and Analytical Chemistry, NMR Group, Institute for Materials Research (Imo-Imomec), Hasselt University, Agoralaan 1, B-3590 Diepenbeek, Belgium; elien.derveaux@uhasselt.be
3   Department of Respiratory Medicine, Ziekenhuis Oost-Limburg, Synaps Park 1, B-3600 Genk, Belgium; michiel.thomeer@zol.be
4   Department of Respiratory Medicine, University Hospital Leuven, Herestraat 49, B-3000 Leuven, Belgium; evelyne.louis@uzleuven.be
5   Department of Respiratory Medicine, Algemeen Ziekenhuis Vesalius, Hazelereik 51, B-3700 Tongeren, Belgium
6   Department of Nuclear Medicine, Ziekenhuis Oost-Limburg, Synaps Park 1, B-3600 Genk, Belgium
*   Correspondence: peter.adriaensens@uhasselt.be

Abstract: Lung cancer is the leading cause of cancer-related mortality worldwide. In order to improve its overall survival, early diagnosis is required. Since current screening methods still face some pitfalls, such as high false positive rates for low-dose computed tomography, researchers are still looking for early biomarkers to complement existing screening techniques in order to provide a safe, faster, and more accurate diagnosis. Biomarkers are biological molecules found in body fluids, such as plasma, that can be used to diagnose a condition or disease. Metabolomics has already been shown to be a powerful tool in the search for cancer biomarkers since cancer cells are characterized by impaired metabolism, resulting in an adapted plasma metabolite profile. The metabolite profile can be determined using nuclear magnetic resonance, or NMR. Although metabolomics and NMR metabolite profiling of blood plasma are still under investigation, there is already evidence for its potential for early-stage lung cancer diagnosis, therapy response, and follow-up monitoring. This review highlights some key breakthroughs in this research field, where the most significant biomarkers will be discussed in relation to their metabolic pathways and in light of the altered cancer metabolism.

Keywords: biomarkers; lung cancer; metabolomics; NMR (nuclear magnetic resonance)

1. Introduction

Lung cancer is the leading cause of cancer-related mortality worldwide, with an overall five-year survival rate varying from 4 to 17% [1]. These low survival rates are mainly caused by delayed diagnosis that often occurs at an advanced stage of the disease due to the absence of symptoms in the early stages. Early detection of lung cancer is, however, of utter importance to improve survival rates [2]. One of the most promising techniques for early detection is low-dose computed tomography (LDCT) [3]. Studies have already shown that LDCT reduces lung cancer mortality by 20% compared to radiography [4]. The NELSON trial reported an even higher reduction of 24–33%, depending on gender, compared to no screening [5]. Nevertheless, this technique does not come without its pitfalls. One of the main challenges of LDCT is the distinction between benign and malignant lesions, resulting in overdiagnosis and high false-positive rates in individuals without lung cancer [6]. The data obtained in the National Lung Screening Trial (NLST) also revealed some additional concerns: in order to reach the desired benefit, i.e., a successful case of

early lung cancer detection in a high-risk population, at least 320 high-risk individuals should be screened for three rounds [7]. To achieve this, many patients have to be exposed to the emotional stress and physical risks associated with screening, such as radiation carcinogenesis. Another important technique for lung cancer detection is positron emission tomography (PET), in which quantitative images of regional in vivo biology are obtained by means of 18F-fluorodeoxyglucose (^{18}F-FDG) tracer uptake [8]. This technique is characterized by a high sensitivity of 96%, indicating that almost no lung lesions remain undetected [9]. Nevertheless, an increased ^{18}F-FDG uptake is not directly related to cancer. Several benign PET-positive lesions, such as inflammation or tuberculosis, are also characterized by an increased ^{18}F-FDG signal. Therefore, misdiagnosis and false-positive results arise as a result of misinterpretation [10]. As a result, about 10% of the patients with non-metastasized PET-positive lung lesions underwent unnecessary surgical interventions [11]. Many efforts have been made to solve these issues, such as changing tracers or measuring protocols, but all alternatives were found to be less appropriate in daily clinical practice [12].

Due to all these challenges, researchers are still looking for early biomarkers that can be used to complement existing screening techniques to provide a safe, faster, and more accurate diagnosis. Biomarkers are defined as "characteristics that are objectively measured and evaluated as indicators of normal biological processes, pathological processes, or pharmacological responses to therapeutic interventions" [13] (p. 344). A new platform that has evolved over recent years for identifying potential early-stage biomarkers is metabolomics, which is defined as the analysis of a large number of metabolites in biological fluids [14]. Since cancer cells are rapidly growing, and therefore have an increased energy consumption, metabolic reprogramming occurs as soon as the disease arises. These metabolic changes eventually result in altered metabolite concentrations in the plasma of cancer patients compared to those found in the plasma of healthy controls [15].

Disruptions are found in several metabolic pathways including the tricarboxylic acid (TCA) cycle, glycolysis, oxidative phosphorylation (OXPHOS), and lipid and amino acid metabolism [16]. One of the most discussed metabolic changes that occurs within cancer cells is the 'Warburg effect', which was originally characterized by an increased secretion of lactate even in the presence of normal oxygen concentrations [17]. This phenomenon was attributed to a defective OXPHOS or mitochondrial function, assuming that the mitochondrial metabolism only contributes little to energy production. More recent research has however shown that most cancer cells have functional mitochondria, allowing for the upregulation of both glycolysis and OXPHOS in order to fulfill the high anabolic demands of cancer cells [15]. Another fundamental metabolite that supports cancer cell proliferation is the most abundant amino acid in human blood: glutamine. Glutamine serves as an alternative carbon source for energy and building block production [15]. In the case of mitochondrial dysfunction, glutamine also plays an important role as its oxidation and carboxylation replenish the TCA cycle, allowing for further biosynthetic reactions to take place, such as the formation of fatty acids. Malignant cells are characterized by an upregulation of components of fatty acid synthesis, used for the formation of lipid bilayers and an increased amount of saturated fatty acids in the cell membrane to resist oxidative damage [15]. All of the above-mentioned processes ultimately lead to altered metabolite concentrations, resulting in a unique metabolite profile.

One of the main tools that is currently used for metabolic profiling, besides mass spectrometry, is proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy since the ^{1}H nucleus is omnipresent in metabolites [16]. ^{1}H-NMR has several advantages over mass spectrometry such as high reproducibility (>98%), short measuring time, and simple sample pre-treatment. As the detected peak areas correlate directly with the metabolite concentration, this technique also allows for the accurate quantitative analysis of substances even in complex mixtures such as plasma [16]. In this review, several key breakthroughs in the search for early-stage biomarkers in the plasma of lung cancer patients by ^{1}H-NMR will be discussed, and the altered metabolic pathways of these biomarkers will be discussed in

the context of lung cancer. Additionally, the potential of these metabolites during screening and diagnosis, as well as their functionality during prediction and follow-up, will be shown.

This review presents the evolution of metabolomics studies in which NMR spectroscopy was used to discover metabolic biomarkers in the plasma of lung cancer patients. This evolution is first described systematically by the studies presented in parts I, II, and III. Then, metabolites found to serve as biomarkers for early-stage NSCLC will be discussed, along with their biochemical pathways in the context of the onset of impaired cancer metabolism.

In part I, the feasibility of lung cancer detection via blood plasma using NMR spectroscopy is demonstrated. First, a study by Louis et al. is discussed, in which lung cancer patients were distinguished from healthy controls based on their metabolite plasma profiles [18]. Based on this study, a second study was performed in which lung cancer plasma profiles were compared with the cancer plasma profiles of breast cancer [19]. These studies demonstrate the feasibility of using NMR spectroscopy to use plasma biomarkers to diagnose lung cancer and to differentiate between different cancer types.

In part II, Derveaux et al. used a higher field strength NMR (600 MHz instead of 400 MHz) and the addition of a human serum albumin-binding (HSA-binding) competitor to further improve resolution and sensitivity and to avoid the binding of specific metabolites to HSA, respectively [20]. In this study, the metabolite plasma profiles of lung cancer patients and healthy controls were again compared, but now with the addition of TSP as an HSA-binding competitor to the sample. This is crucial to prevent the binding of plasma metabolites to HSA, an abundant protein in human plasma, and thus allows for a more accurate determination of the plasma metabolite concentrations. Another important breakthrough shows that, in the presence of TSP, maleic acid (MA) can be used as an internal standard. While peak area normalization is often performed in the total spectral area of the sum of all metabolites, the single signal of MA can now be used. The combination of a stronger magnetic field and an improved sample preparation (i.e., the addition of TSP as an HSA-binding competitor and MA as an internal standard) thereby allows for a more accurate analysis of the altered metabolism in cancer.

Part III highlights a second study by Derveaux et al. that focuses on a deeper metabolic interpretation of early-stage lung cancer by comparing pre- and postoperative plasma metabolite profiles of patients diagnosed with early-staged NSCLC [21]. Importantly, pre- and postoperative plasma samples were donated by the same patients at different timepoints before and after complete surgical resection (i.e., lobectomy). Collection of plasma samples via this longitudinal study design leads to the reduction in patient-to-patient variation and the specification of the plasma metabolic phenotype for early-stage lung cancer. This study shows the potential of plasma biomarkers to (i) monitor therapy/surgery response and (ii) detect early disease recurrence in lung cancer.

A graphical overview of these three parts is presented as Figure A1 in Appendix A.

2. Evolution in NMR Metabolomics Research—Part I: Feasibility of Lung Cancer Detection via Blood Plasma Using NMR Spectroscopy

2.1. Feasibility: Lung Cancer Detection

2.1.1. Study Set-Up

The first study by Louis et al. investigated whether lung cancer patients could be distinguished from healthy controls based on the plasma metabolic phenotype with the use of ^{1}H-NMR [18]. A total of 357 lung cancer patients and 347 controls, defined as patients with noncancerous disease, were included. If all inclusion criteria were fulfilled, fasting venous blood samples were collected and processed to obtain plasma aliquots for ^{1}H-NMR analysis. The obtained NMR spectrum was segmented into 110 spectral regions based on metabolite spiking, in which a small quantity of known metabolites was added to the sample to identify the peak positions [19]. This was essential since metabolomics was a relatively new discipline, with a lack of reference NMR spectra for all known metabolites [22]. This method is preferred to obtain an optimal signal assignment certainty. These 110 normalized integration regions were the variables used for multivariate

statistics. The integration regions are the variables for the multivariate statistics, and the value of each variable is determined by the area under the peak of the integration region (and therefore directly correlated with the concentration of a specific metabolite). The NMR measurements were carried out using a 400 MHz NMR spectrometer. More technical details can be found in Appendix B or in the paper of Louis et al. [18].

For the statistical analysis, both lung cancer patients and controls were divided into a training and an independent validation cohort. Subdivision into these two groups is of great importance since external validation is essential to provide reliable conclusions and to confirm, or refute, the strength of the statistical model. Prior to this external validation, a 7-fold internal cross-validation was carried out in the training cohort where validation occurs repeatedly using different subsets of the same dataset. Before subdivision occurred, a principal component analysis (PCA) was carried out to identify confounders. Some outliers were excluded based on Hotelling's T2 range test, resulting in 331 lung cancer patients and 315 controls. From these cohorts, 233 lung cancer patients and 226 controls were assigned to the training cohort, leaving 98 lung cancer patients and 89 controls for the validation cohort (Table 1). The most common methods used for multivariate statistics include PCA, partial least squares projection to latent structures (PLS), orthogonal projection to latent structures (OPLS), and orthogonal partial least squares discriminant analysis (OPLS-DA). OPLS-DA was used in the studies described here, which is common in NMR-metabolomics. The obtained statistical model can then be interpreted using different model parameters: R2X (the goodness of fit; variation within the groups explained by the model; indicated on the predictive vertical axis), R2Y (explanation between the groups; indicated on the horizontal axis), and Q2 (the predictive accuracy of the trained model) [21]. The closer the values of R2X and R2Y to 1, the better the model. A Q2 value > 0.30 is considered a highly predictive model in the context of metabolomics [21]. Training of the classification model is always performed using data from the training cohort, while additional independent validation of the model is always performed using data from the independent validation cohort, where no group information is provided. High sensitivity and specificity rates in this validation cohort will, therefore, indicate the actual differentiation between the two groups [21]. The variables responsible for the observed differentiation can be analyzed using the variable of importance of projection (VIP) analysis. Another important parameter that is taken into account is the area under the curve (AUC) value of the receiver operating curves (ROCs), which indicates the overall performance of both training and validation models. It is assumed that the closer the AUC value to 1, the better the model. More information about OPLS-DA analysis can be found in Appendix B or in the book of Eriksson et al. [23].

Table 1. Characteristics of the study and statistical analysis.

Louis et al., 2016 [18]				
	Training Cohort		Validation Cohort	
	LC	C	LC	C
Number of subjects, n	233	226	98	89
Sensitivity (%)	78		71	
Specificity (%)	92		81	
R2X (Cum)	0.864		-	
R2Y (Cum)	0.477		-	
Q2 (Cum)	0.391		-	
AUC	0.88		0.84	

AUC, area under the curve; C, controls; LC, lung cancer patients; R2X (Cum), total explained variation in X within groups; R2Y (Cum), total explained variation in Y between groups; Q2 (Cum), predicted variation.

2.1.2. Results

As mentioned above, an OPLS-DA model was first trained using the training cohort (Table 1). This resulted in a model that was able to distinguish the metabolite plasma profiles of the lung cancer patients from those of the controls with a sensitivity of 78% and a specificity of 92% (Figure 1a) (Table 1). Since the trained OPLS-DA model showed an R2X and R2Y of 0.864 and 0.477, respectively, and a Q2 value of 0.391, a good overall quality and predictive accuracy can be assumed (Table 1). The statistical model was then applied to the independent validation cohort, showing a sensitivity and specificity of 71% and 81%, respectively (Figure 1b) (Table 1). The overall performance of the training and validation models was confirmed by AUC values of 0.88 and 0.84, respectively (Figure 1c) (Table 1).
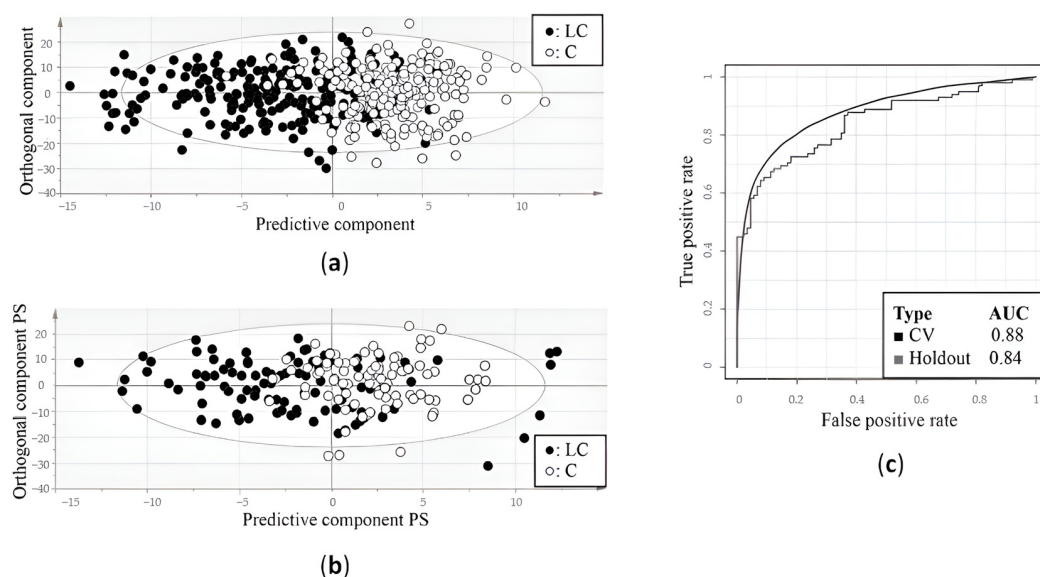
Figure 1. Differentiation between lung cancer patients and healthy controls based on metabolic profiling in plasma. (**a**) Orthogonal partial least squares discriminant analysis (OPLS-DA) score plot from the training cohort shows a distinction between the plasma metabolite profiles of lung cancer patients and controls with a sensitivity of 78% and a specificity of 92%. (**b**) OPLS-DA score plot from the independent validation cohort shows a distinction between the metabolite profiles of lung cancer patients and controls with a sensitivity of 71% and specificity of 81%. (**c**) Receiver operating curves show the high predictive accuracy of the OPLS-DA model of the training cohort for both cross-validation and independent validation, reaching AUC values of 0.88 and 0.84, respectively. AUC, area under the curve; C, controls; CV, cross-validation; LC, lung cancer patients; PS, predicted scores. Figure adapted from reference [18].

2.1.3. Supporting Evidence

Puchades-Carrasco et al. analyzed whether serum metabolic profiling could be used to diagnose lung cancer [24]. Serum samples were collected from 182 lung cancer patients (both early and advanced stages) and 87 healthy controls. First, PCA was performed to evaluate intrinsic variability within the data, clustering trends, and potential outliers. Afterward, OPLS-DA was carried out using a training cohort of 142 lung cancer patients and 74 controls, and a validation cohort of 40 lung cancer patients and 13 controls. This study showed that serum metabolic profiling allowed for the differentiation between lung cancer patients and healthy controls with a sensitivity and specificity of 92.3% and 95%.

Rocha et al. investigated the plasma metabolite profiles of 85 primary lung cancer patients and 78 healthy controls [25]. PCA analysis was first performed to detect clusters and outliers within the dataset, followed by OPLS-DA. Monte Carlo Cross Validation (MCCV) was used to check the predictive ability of the model by repeatedly using different subsets of the same dataset [26]. Some disadvantages of this method include its computational intensity, the risk of overfitting, and subsampling bias which jeopardizes the model's validity

and reliability [27]. In the study, a sensitivity and specificity of 92% and 89% were reached, together with a Q2 score of 0.64. The results confirm the potential of plasma metabolite profiling for lung cancer diagnosis and screening.

A study by Carrola et al. further supported these findings using urine instead of plasma [28]. Urine samples from 71 lung cancer patients and 54 healthy controls were compared. PCA was applied to detect clusters and outliers, followed by OPLS-DA and MCCV analyses. Samples from lung cancer patients and healthy controls could be distinguished, with a sensitivity and specificity of 93% and 94%, respectively, together with an overall classification rate of 93.5%.

2.2. Feasibility: Differentiation between Cancer Types: Lung Cancer versus Breast Cancer

Previous experiments demonstrated the feasibility of plasma metabolite alterations to differentiate lung cancer patients from healthy controls. This raised the question of whether the metabolic fingerprint was a general cancer marker or whether it could also be used to differentiate between different types of cancer. In a second study by Louis et al., it was demonstrated that the plasma metabolite profile allows one to differentiate between lung and breast cancer [19].

2.2.1. Study Set-Up

A total of 145 lung cancer patients with adenocarcinoma and 147 breast cancer patients with adenocarcinoma were enrolled in the study. If all inclusion criteria were fulfilled, fasting venous blood samples were collected and processed in order to obtain plasma aliquots used for ^{1}H-NMR analysis. The NMR analysis is as described above in Section 2.1.1, resulting in 110 integration values for the statistical analyses. More technical details can be found in Appendix B or in the paper of Louis et al. [19].

After removing outliers (Hotelling's T2 test), a total of 135 lung cancer patients and 140 breast cancer patients were subdivided into a training and validation cohort. The training cohort consisted of 54 female lung cancer patients and 80 breast cancer patients, leaving 81 male lung cancer patients and 60 breast cancer patients for the validation cohort (Table 2). Male and female lung cancer patients were allocated in separate groups to exclude a confounding effect of gender.

Table 2. Characteristics of the study and statistical analysis.

Louis et al., 2016 [19]				
	Training Cohort		**Validation Cohort**	
	LC	**BC**	**LC**	**BC**
Number of subjects, n	54	80	81	60
Sensitivity (%)	93		89	
Specificity (%)	99		82	
R2X (Cum)	0.82		-	
R2Y (Cum)	0.73		-	
Q2 (Cum)	0.63		-	
AUC	0.96		0.94	

AUC, area under the curve; C, controls; LC, lung cancer patients; R2X (Cum), total explained variation in X within groups; R2Y (Cum), total explained variation in Y between groups; Q2 (Cum), predicted variation.

2.2.2. Results

The OPLS-DA classification model was first trained using the training cohort. This resulted in a sensitivity of 93% (93% of the lung cancer patients were correctly classified) and a specificity of 99% (99% of the breast cancer patients were correctly identified), indicating that almost all female breast cancer patients were correctly identified. This was further confirmed by an AUC of 0.96 (Table 2) (Figure 2a,b). The predictive accuracy was then determined using the independent validation cohort, showing a sensitivity of 89% and specificity of 82%, together with an AUC of 0.94 (Figure 2b,c).
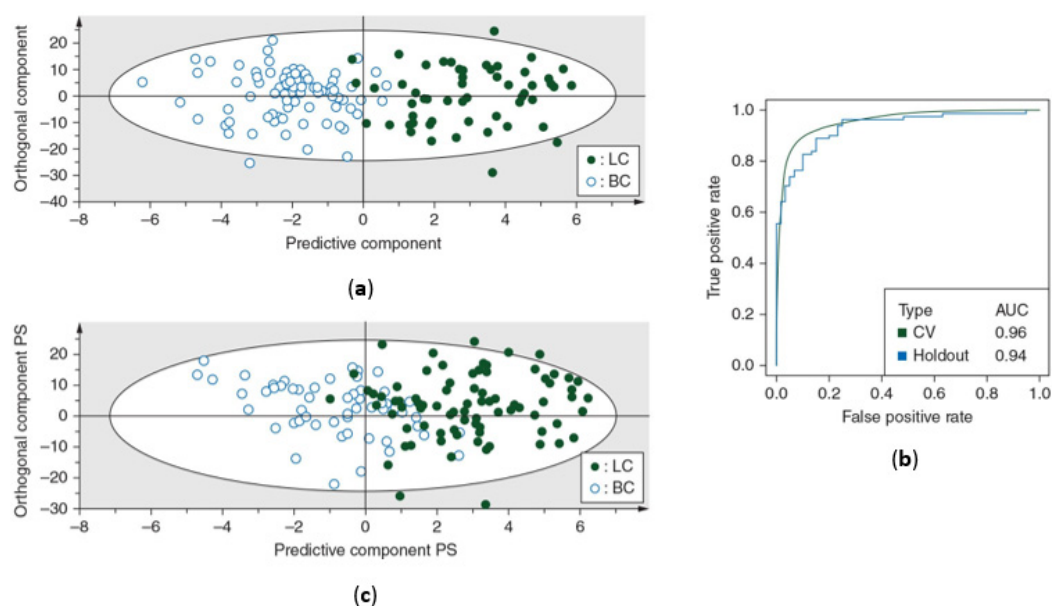
Figure 2. Differentiation between lung and breast cancer patients based on metabolic profiling in plasma. (**a**) Orthogonal partial least squares discriminant analysis (OPLS-DA) score plot from the training cohort shows a distinction between the metabolite profiles of lung cancer patients and breast cancer patients with a sensitivity of 93% (93% of the lung cancer patients were correctly classified) and a specificity of 99% (99% of the breast cancer patients were correctly classified). (**b**) Receiver operating curves show the high predictive accuracy of the OPLS-DA model of the training cohort for both cross-validation and independent validation, reaching AUC values of 0.96 and 0.94, respectively. (**c**) OPLS-DA score plot from the independent validation cohort shows a distinction between the metabolite profiles of lung cancer patients and breast cancer patients with a sensitivity of 89% and specificity of 82%. AUC, area under the curve; BC, breast cancer patients; CV, cross-validation; LC, lung cancer patients; PS, predicted scores. Figure adapted from reference [19].

Another study by Louis et al. further confirmed the ability of the metabolic fingerprint to differentiate between different cancer types, where breast, colon, and lung cancer were successfully differentiated [29]. Here, 95% of the breast cancer patients, 78% of the colorectal cancer patients, and 84% of the lung cancer patients were correctly identified.

3. Evolution in NMR Metabolomics Research—Part II: Developments in Preanalytical Sample Preparation and NMR Measurement Procedure

The studies discussed above have shown the feasibility of plasma metabolic biomarkers as a tool for screening and diagnosis of lung cancer, as well as for the differentiation between different cancer types. These studies were all performed using a 400 MHz NMR spectrometer. Since these experiments were feasible, i.e., they were able to distinguish the lung cancer patients from controls based on plasma metabolic phenotypes, it was worth investing in a stronger magnet of 600 MHz with higher sensitivity. The higher resolution allows for the definition of more specific integration regions, i.e., more integration regions that contain a signal representing one single metabolite, eventually resulting in a more accurate statistical analysis and identification of the most differentiating metabolites. While distinguishing plasma metabolite profiles of lung cancer and controls was initially the primary focus, the use of a stronger magnetic field could contribute more to a better understanding of impaired cancer metabolism.

3.1. TSP as an HSA Binding Competitor and MA as an Internal Standard

Although several metabolomics studies have identified alterations in metabolite concentrations within the plasma of lung cancer patients, an essential factor has often been overlooked. Derveaux et al. demonstrated that, since several plasma metabolites

have a high affinity for the plasma protein HSA, only the HSA-unbound fraction of these metabolites is measured [20]. In other words, the ^{1}H-NMR signal intensities of these metabolites do not reflect their total concentration, and this will, therefore, lead to an underestimation of the total plasma metabolite concentration [20]. In addition, this study showed that some metabolites are not only characterized by this affinity towards HSA but that the commonly used internal (chemical shift) reference TSP also strongly binds to HSA. Therefore, TSP cannot be used as an internal reference since the obtained signal no longer reflects the amount added during the sample preparation [30].

However, TSP has another major advantage that has not been described before. Since TSP has a much higher affinity for HSA than the other plasma metabolites, its addition leads to the saturation of the HSA metabolite-binding sites and thus to the dissociation of these metabolites from HSA (Figure 3). TSP could, therefore, be used as an HSA-binding competitor: by preventing the binding of other metabolites to HSA, the total metabolite plasma concentrations can be measured. It can, therefore, be concluded that TSP addition is crucial for the accurate determination of total plasma metabolite concentrations in metabolomics studies.
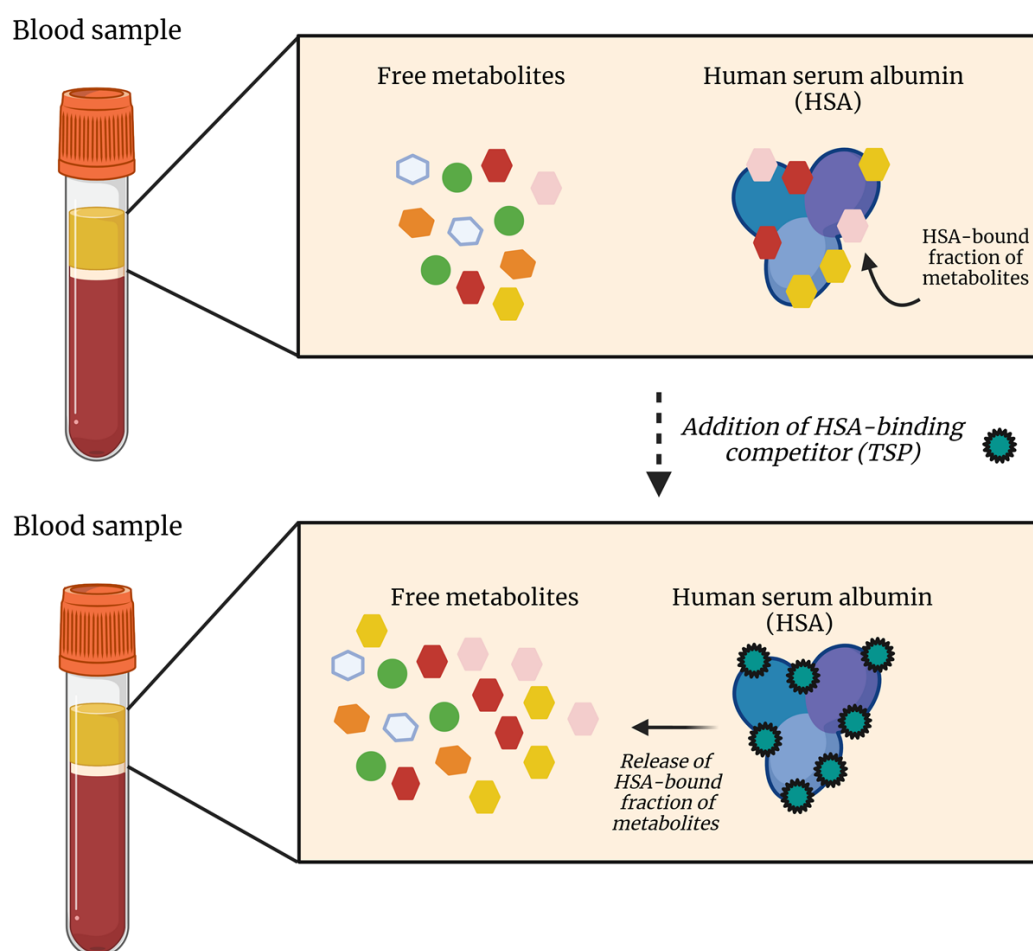
Figure 3. The addition of TSP leads to the dissociation of the HSA-bound fraction of metabolites due to the extremely high affinity of TSP for HSA. While some plasma metabolites do not show affinity for HSA, others are attracted and bind to HSA, resulting in an underestimation of the actual plasma metabolite concentration. After adding TSP, which is characterized by an even higher HSA affinity compared to the plasma metabolites, TSP will bind to HSA, resulting in the release of the bound metabolites. HSA, human serum albumin; TSP, trimethylsilyl-2,2,3,3-tetradeuteropropionic acid. Figure created with BioRender.com.

Additionally, due to its high affinity for HSA, TSP is not a suitable candidate as the internal standard for quantification, indicating that another internal standard should be used in order to quantify plasma metabolites. Derveaux et al. proposed MA as a suitable candidate for absolute quantification and peak area normalization since it gives rise to a sharp, isolated signal (i.e., around 6 ppm in the ^{1}H NMR spectrum) without overlapping with other metabolite signals [20]. In addition, MA can be purchased with excellent purity, it is easily dried, and it has a distinct solubility in D_2O [20]. To make sure that MA could be used as an internal standard, the affinity towards HSA was analyzed. Here, it was shown that MA showed some affinity towards HSA but, as for several other metabolites, at a lower degree compared to TSP. Therefore, when added together with TSP, MA did not bind to HSA, enabling it to be used as a reliable internal standard for quantification. Knowing this, the plasma metabolite profiles of lung cancer patients were compared with those of controls in the presence of 4 mM TSP as an HSA-binding competitor and MA as the internal standard.

3.2. Study Design

In order to validate the new proposed methodology, a total of 141 lung cancer patients and 135 healthy controls were included. If all inclusion criteria were fulfilled, fasting venous blood samples were collected as described before and processed to obtain plasma aliquots for ^{1}H-NMR analysis. This led to a total of 114 lung cancer patients and 118 healthy controls for further analysis. In contrast to previous studies, the samples were enriched with the addition of 4 mM TSP as the HSA-binding competitor and 53.58 μM MA as the internal standard. Similarly to the previous studies, metabolite spiking was performed to divide the NMR spectrum into well-defined integration regions. Metabolite spiking with 62 metabolites resulted in the definition of 237 integration regions in the ^{1}H-NMR spectrum recorded on a 600 MHz spectrometer [20]. These 237 normalized integration regions were the variables used for multivariate statistics. More technical details can be found in Appendix B or in the paper of Derveaux et al. [20].

Both patient groups were subdivided into a training and validation cohort, where the training cohort consisted of 80 lung cancer patients and 80 controls, and the validation cohort consisted of 34 lung cancer patients and 38 controls (Table 3).

Table 3. Characteristics of the study and statistical analysis.

Derveaux et al., 2021 [20]				
	Training Cohort		**Validation Cohort**	
	LC	**C**	**LC**	**C**
Number of subjects, n	80	80	34	38
Sensitivity (%)	85		74	
Specificity (%)	93		74	
R2X (Cum)	0.861		-	
R2Y (Cum)	0.581		-	
Q2 (Cum)	0.364		-	
AUC	0.95		-	

AUC, area under the curve; C, controls; LC, lung cancer patients; R2X (Cum), total explained variation in X within groups; R2Y (Cum), total explained variation in Y between groups; Q2 (Cum), predicted variation.

3.3. Results

Data reduction was performed (cross-validation; 'jack-knifing'), leaving 78 variables for final statistical analysis. Using these 78 variables, an unsupervised PCA plot was created to gain some first insights into the separation of the data. A clear distinctive trend was observed between the plasma metabolite profiles of the lung cancer patients

and controls within the training cohort (Figure 4a). Afterward, the same 78 variables were used to train a supervised OPLS-DA model. The constructed model was able to distinguish the two groups with a sensitivity of 85% and a specificity of 93% (Table 3) (Figure 4b). The strength of this model was further confirmed by high values of the R2X (Cum), R2Y (Cum), and Q2 (Cum) parameters of 0.861, 0.581, and 0.36, respectively (Table 3). The overall performance was also assessed by the ROC curve, where an AUC value of 0.95 was obtained (Table 3) (Figure 4c). Afterward, the obtained results were validated via the independent validation cohort. The ability to distinguish the two groups remained, with a sensitivity and specificity of 74% each (Table 3) (Figure 4d). The proposed methodology allowed for (i) the measurement of actual metabolite concentrations using TSP and (ii) the quantification of absolute plasma metabolite and normalization of peak area using MA. This enabled a precise study of the plasma metabolite profile in order to identify potential biomarkers.
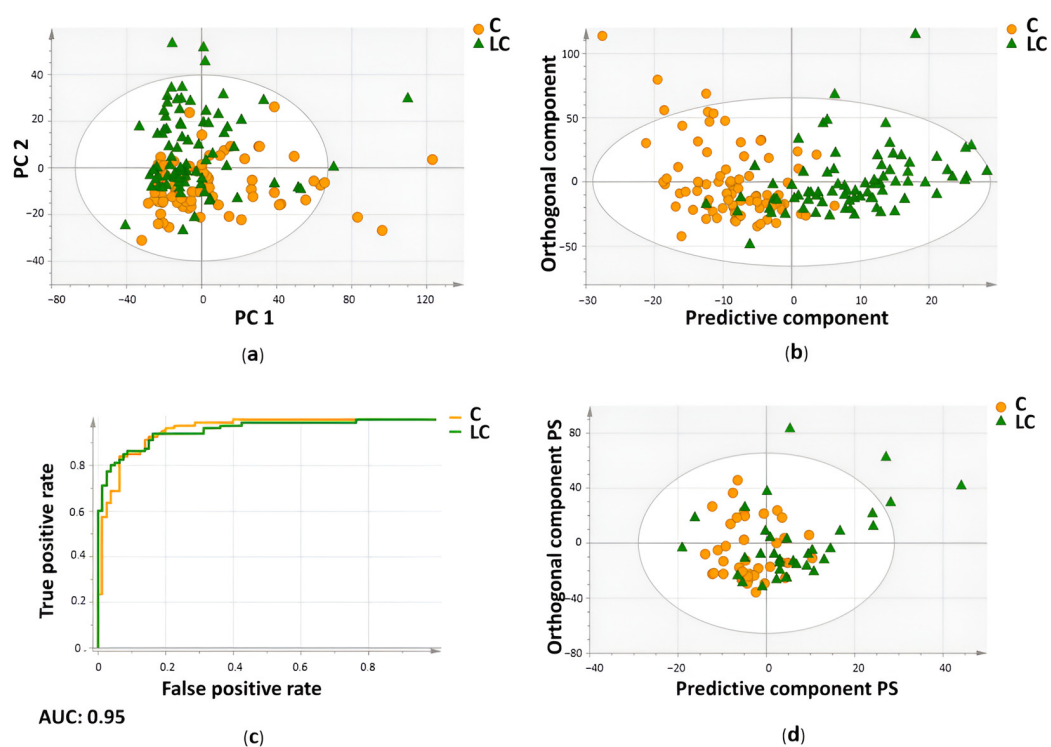
Figure 4. Differentiation between lung cancer patients and healthy controls based on metabolic profiling in plasma using a new methodology including TSP, as a human serum albumin (HSA)-binding competitor, and MA as an internal standard for metabolite quantification. (**a**) Unsupervised principal component analysis (PCA) shows a distinctive trend between the plasma metabolite profiles of the lung cancer patients (green triangles) and controls (orange circles). (**b**) Orthogonal partial least squares discriminant analysis (OPLS-DA) shows a clear distinction between the plasma metabolite profiles of the 80 lung cancer patients and 80 controls within the training cohort. The constructed model was characterized by a sensitivity and specificity of 85% and 93%, respectively. (**c**) The receiver operating characteristic (ROC) curve of the OPLS-training model confirmed the overall performance with an AUC of 0.95. (**d**) Validation of the constructed OPLS-DA model on an independent validation cohort resulted in a distinction between the plasma metabolite profiles of the 34 lung cancer patients and 38 controls with a sensitivity and specificity of 74% each. AUC, area under the curve; C, controls; LC, lung cancer patients; MA, maleic acid; PC, principal component; PS, predicted scores; TSP, trimethylsilyl-2,2,3,3-tetradeuteropropionic acid. Figure adapted from reference [20].

Therefore, further analysis was performed to identify the most significant contributors to the observed discrimination between the two groups. The most significant decreases in the plasma of lung cancer patients included lipid signals from fatty acid chains, phosphatidylcholines, and sphingomyelins. The most significant increases in the plasma of lung cancer patients included glucose, isoleucine, glycerol, and isopropanol [20]. The decrease in the above-mentioned lipids has been widely discussed in literature and is attributed to various factors including enhanced membrane synthesis, increased enzyme activity, and precursor consumption in lung cancer.

4. Evolution in NMR Metabolomics Research—Part III: Plasma Biomarkers for Early-Stage NSCLC and Their Potential for Detection and Monitoring of NSCLC Recurrence

Derveaux et al. improved metabolite quantification and the metabolite fingerprinting of lung cancer by adding a more accurate internal standard (MA) in combination with an HSA-binding competitor TSP and using a stronger NMR magnetic field strength [20]. However, this can even be further improved by reducing patient-to-patient (donor-intrinsic) variabilities and modeling only the metabolic changes that are related to the onset of lung cancer. To overrule these variabilities, Derveaux et al. performed another study with a longitudinal setup in which the pre- and postoperative plasma metabolite profiles of the same individuals, all early-stage NSCLC patients who received surgery, were compared [21]. Analysis of the pre- and postoperative plasma samples provides early-stage NSCLC and healthy metabolite profiles, respectively. As each patient can now serve as its own control, paired multivariate statistics can be used to differentiate pre-and postoperative plasma metabolite profiles where patient-to-patient variations are reduced. The most differentiating plasma metabolites in these models are, therefore, very specific biomarkers for early-stage NSCLC. This unique metabolomics study shows the potential of early detection of NSCLC recurrence by non-invasively monitoring NSCLC patients during therapy follow-up via blood samples.

4.1. Study Design

A total of 114 lung cancer patients diagnosed with resectable stage I-IIIA non-small cell lung cancer (NSCLC) were included. In case all inclusion criteria were fulfilled, fasting venous blood samples were collected at three different time points, including baseline (B; day of surgery; preoperative), control (C; after diagnosis but before surgery; preoperative), and effect (E; three months after surgery; postoperative-3M). For the effect group, it was clinically confirmed that none of these patients showed disease recurrence after three months. All blood samples were processed to obtain plasma aliquots for ^{1}H-NMR analysis. The same NMR procedure was applied as proposed in the previous study of Derveaux et al., where (i) a 600 MHz NMR spectrometer was used, (ii) TSP was added as an HSA-binding competitor, and (iii) MA was added as a quantitative internal standard [20]. Again, the NMR spectrum was divided into 237 well-defined integration regions which were used as variables for further statistical analyses. In case (i) no complete surgical resection could be performed, (ii) a postoperative pathological non-malignant diagnosis appeared, or (iii) disease recurrence occurred within a period of 6 months, patients were excluded from the study. This resulted in a total of 74 remaining early-stage lung cancer patients (Table 4) [21]. More technical details can be found in Appendix B or in the paper by Derveaux et al. [21].

All patients were subdivided into a training cohort (50 patients) and an independent validation cohort (24 patients) (Table 4). Since three different comparisons were aimed for (B vs. E; C vs. E; B vs. C), three different statistical models were constructed. The third comparison, B vs. C, was carried out as an additional control to find out whether the metabolite profiles at these two preoperative time points are similar and patient-specific.

Table 4. Characteristics of the study and statistical analysis.

Derveaux et al., 2023 [21]					
	B/E		**C/E**		**B/E**
	OPLS-DA	**OPLS-EP**	**OPLS-DA**	**OPLS-EP**	**OPLS-DA**
TRAINING COHORT					
Number of subjects, n	50	50	50	50	50
Sensitivity (%)	92	-	88	-	74
Specificity (%)	96	-	90	-	62
R2X (Cum)	0.55	0.59	0.53	0.57	0.31
R2Y (Cum)	0.67	0.89	0.61	0.83	0.15
Q2 (Cum)	0.42	0.76	0.36	0.60	0.08
AUC	0.99	-	0.97	-	0.72
VALIDATION COHORT					
Number of subjects, n	24	24	23	23	23
Sensitivity (%)	88	-	96	-	74
Specificity (%)	92	-	91	-	43
AUC	0.97	-	0.97	-	0.69

Within the control group, one sample was identified as an outlier and excluded from further analyses. AUC, area under the curve; B, baseline (preoperative, before surgery); C, control (preoperative, at time of diagnosis); E, effect (postoperative, 3M after surgery); OPLS-DA, orthogonal partial least squares discriminant analysis; OPLS-EP, orthogonal partial least squares effect projections; R2X (Cum), total explained variation in X within groups; R2Y (Cum), total explained variation in Y between groups; Q2 (Cum), predicted variation.

4.2. Results

A first supervised OPLS-DA model was trained with the data from the baseline (B) and effect (E) timepoints of the training cohort. Here, excellent discrimination between pre- and postoperative profiles was observed with a sensitivity and specificity of 92% (92% of the samples were correctly identified as preoperative) and 96% (96% of the samples were correctly identified as postoperative), respectively (Table 4) (Figure 5a). The predictive accuracy was further strengthened by the high values of the obtained R2X (Cum), R2Y (Cum), and Q2 (Cum) parameters of 0.55, 0.67, and 0.42, respectively (Table 4). These results were then validated in the independent validation cohort, resulting again in an excellent distinction of pre- and postoperative plasma metabolite profiles with a sensitivity and specificity of 88% and 92%, respectively (Table 4) (Figure 5b). Additionally, excellent AUC values of 0.99 and 0.97 were reached within the training and validation cohort, respectively (Table 4) (Figure 6). While OPLS-DA is a very strong method for modeling differentiation between the two groups, this method is not appropriate for dependent, paired samples. Since this study compared plasma samples derived from the same lung cancer patient at different timepoints, OPLS-effect projection (OPLS-EP) studies, which take the paired character of the samples into account, could also be performed. Therefore, an OPLS-EP model was trained and validated as well. Here, even better parameter values of the R2X (Cum), R2Y (Cum), and Q2 (Cum) were found, reaching values of 0.59, 0.89, and 0.76, respectively (Table 4). These values strongly support the differentiation found between pre- and postoperative metabolite profiles. More information about OPLS-EP analysis can be found in Appendix B and in the paper by Jonsson et al. [31].

After comparing B/E timepoints, the second OPLS-DA and OPLS-EP models were trained with the data derived from the control (C) and effect (E) timepoints of the training cohort. Again, the constructed model succeeded in distinguishing pre- and postoperative plasma metabolite profiles with outstanding parameters for both

OPLS-DA and OPLS-EP. OPLS-DA showed a sensitivity and specificity of 88% and 90%, respectively, together with R2X (Cum), R2Y (Cum), and Q2 (Cum) parameters of 0.53, 0.61, and 0.36, respectively (Table 4) (Figure 5c). Validation of the OPLS-DA model resulted in even higher sensitivity and specificity rates of 96% and 91%, respectively (Table 4) (Figure 5d). The excellent performance of the constructed model was confirmed by the high AUC values in both training and validation cohorts of 0.97 each (Table 4) (Figure 6). The OPLS-EP model showed R2X (Cum), R2Y (Cum), and Q2 (Cum) parameters of 0.57, 0.83, and 0.60 (Table 4).
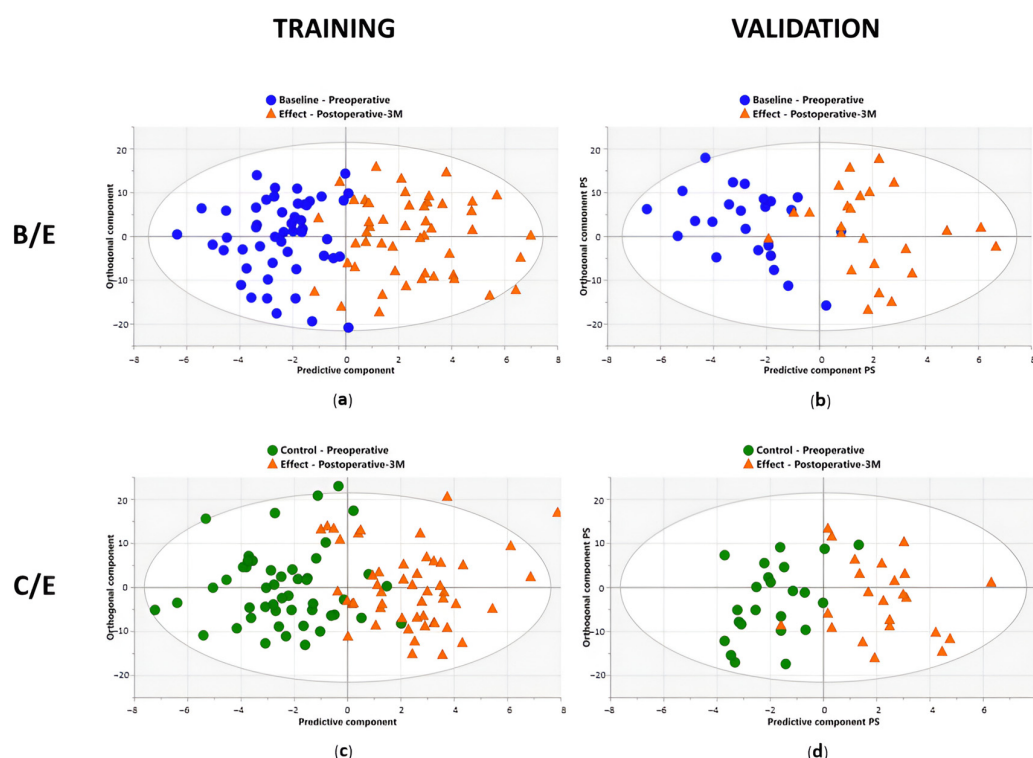
Figure 5. Differentiation between pre- and postoperative plasma samples based on metabolic profiling. The clear separation on the *x*-axis represents the metabolic variation between the pre- and postoperative groups, both when comparing baseline/effect (B/E) and control/effect (C/E). (**a**) Supervised orthogonal partial least squares discriminant analysis (OPLS-DA) shows an excellent distinction between the plasma metabolite profiles of preoperative (baseline: B, blue circles) and postoperative (effect: E (after 3 months), orange triangles) datasets within the training cohort consisting of 50 NSCLC patients. Sensitivity and specificity rates of 92% (92% of the preoperative samples correctly identified as preoperative) and 96% (96% of the postoperative samples correctly identified as postoperative), respectively, were reached. (**b**) Validation of the constructed OPLS-DA model in the validation cohort consisting of 24 NSCLC patients also resulted in a clear distinction between pre- and postoperative profiles with a sensitivity and specificity of 88% and 92%, respectively. (**c**) OPLS-DA model shows a great distinction between the preoperative (control; C; green circles) and postoperative datasets within the training cohort consisting of 50 NSCLC patients. Sensitivity and specificity rates of 88% and 90%, respectively, were reached. (**d**) Validation of the constructed OPLS-DA model in the validation cohort consisting of 23 NSCLC patients again resulted in a clear distinction between pre- and postoperative profiles with a sensitivity and specificity of 96% and 91%, respectively. PS, predictive scores. Figure adapted from reference [21].

As an additional control, a third OPLS-DA model was constructed using data derived from the baseline (B) and control (C). First, the OPLS-DA model was trained using data from the training cohort. This resulted in low R2X (Cum) and R2Y (Cum) values of 0.31 and 0.15, respectively, and a very low Q2 (Cum) of 0.08 (Table 4).
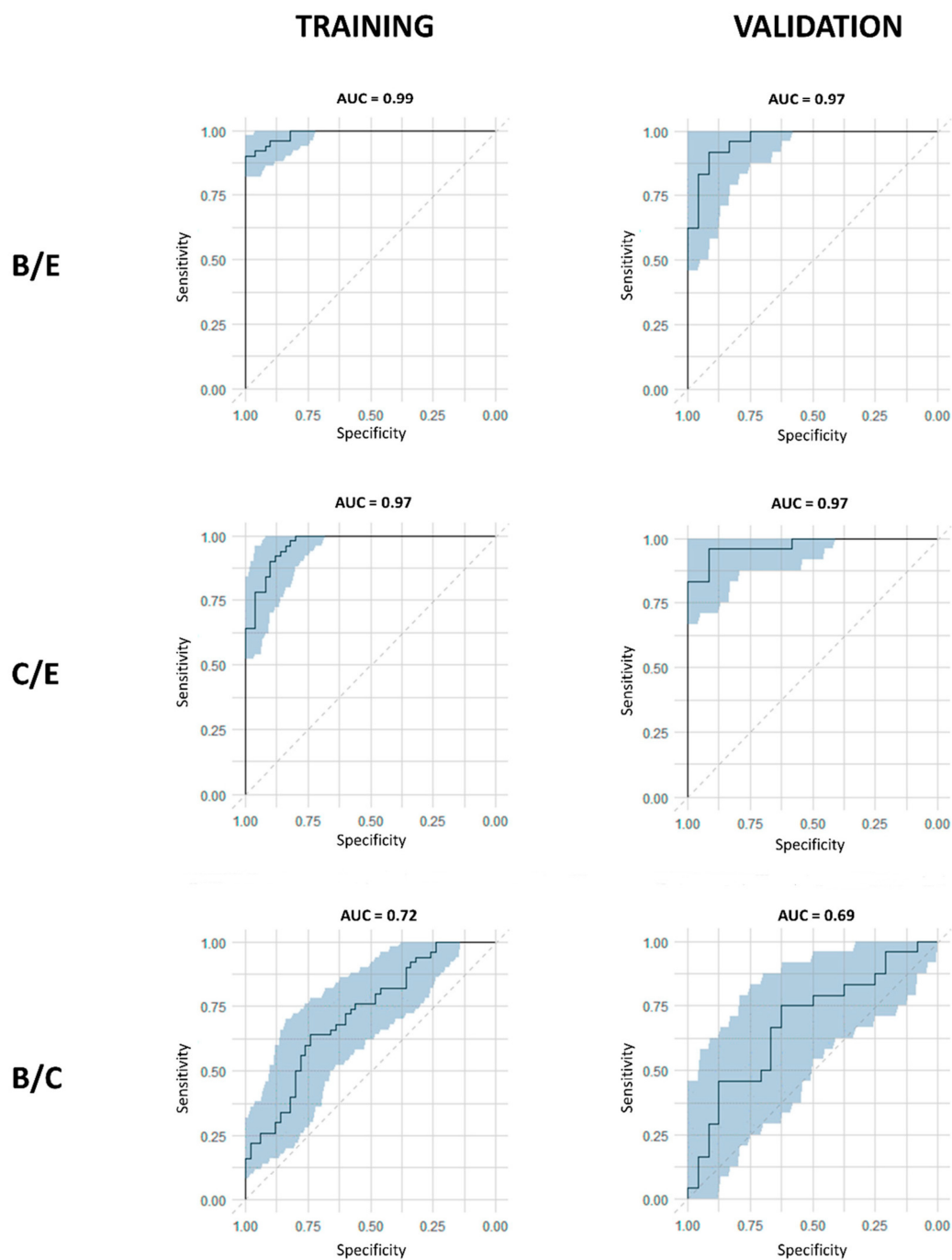
Figure 6. Overview of all the receiver operating curves (ROC), showing the area under the curve (AUC) values, in the three different comparisons, for both training and validation cohorts. Excellent AUC values were reached in both training and validation cohorts when comparing preoperative profiles (baseline: B, and control: C) with postoperative profiles (effect: E). The blue zone represents the 95% confidence interval of the obtained AUC values by internal validation via bootstrapping resampling. Figure reused from reference [21].

In the next step, key contributors to the metabolic shift between the pre- and postoperative profiles were identified. For both the OPLS-DA and OPLS-EP models, the same variables were identified as significant contributors. Preoperative plasma profiles (i.e., lung cancer profiles) showed a decrease in lactate, cysteine, and asparagine levels and an increase in acetate levels. All these changes in concentration were found to be significant (paired *t*-test; $p < 0.001$) (Figure 7) [21].
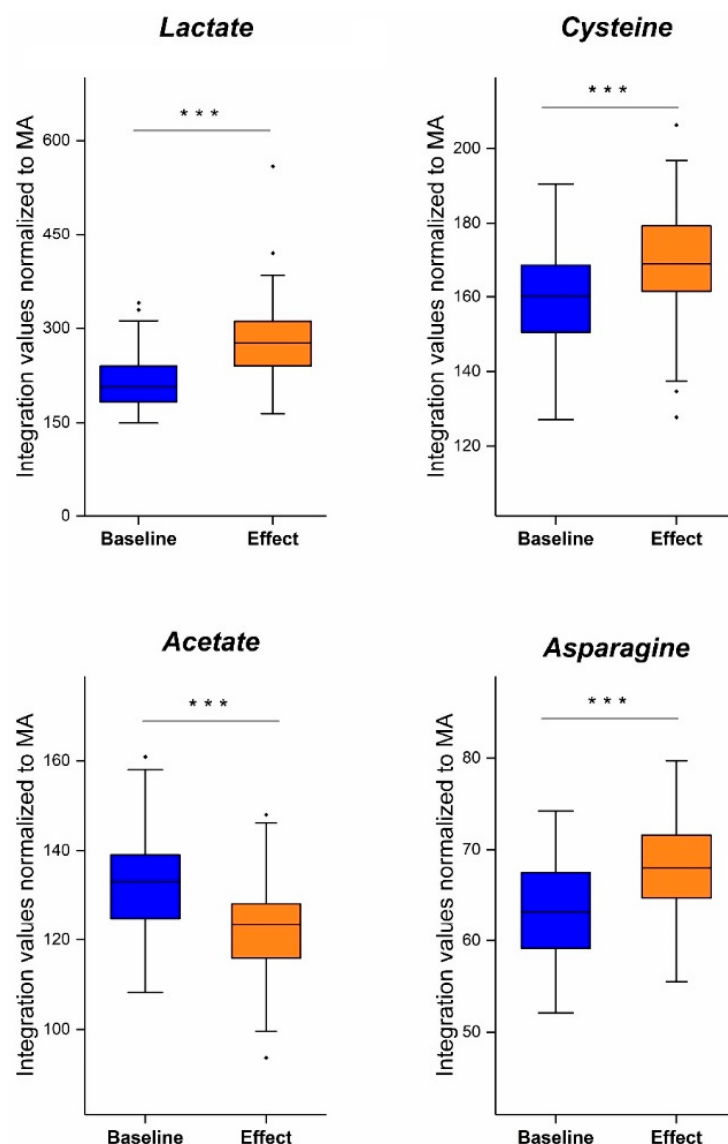
Figure 7. Integration values, normalized to MA, of lactate, cysteine, acetate, and asparagine in preoperative (baseline: blue) and postoperative datasets (effect: orange). Lactate, cysteine, and asparagine showed a significant increase in plasma concentrations after surgery. Acetate showed a significant decrease in plasma concentration after surgery. (*** $p < 0.001$) ($\cdot$ represents a datapoint outside the 1.5 IQR range). MA, maleic acid. Figure reused from reference [21].

5. Discussion of Breakthroughs in NMR Metabolomics Research

It is noteworthy to mention that the models discussed in parts II and III revealed different sets of discriminating key metabolites with increased/decreased plasma levels compared to other metabolomics studies in literature. For example, lactate plasma levels were found to be decreased in the metabolite profile of lung cancer by Derveaux et al., while other studies report increased plasma lactate in lung cancer patients [25,32]. These different outcomes throughout the literature can be explained because of (i) sample preparation with or without the addition of TSP as an HSA-binding competitor, (ii) inclusion of patients diagnosed with different stages of lung cancer, and (iii) patient-to-patient metabolic variability between groups.

The studies discussed in parts II and III involve the addition of TSP to the sample as an HSA-binding competitor. This pre-analytical optimization is of great importance, as some metabolites possess a strong affinity for HSA (e.g., lactate). As a result, without the addition of an HSA-binding competitor, only the unbound fraction of these metabolites is

measured, leading to an underestimation of their total plasma concentrations. Moreover, it is known that the amount of HSA can vary considerably among individuals and time, leading to HSA-dependent differences in measured plasma metabolite levels across different persons and different timepoints.

For example, lactate concentrations can be heavily underestimated without the addition of an HSA-binding competitor and may even lead to the incorrect interpretation of an apparent lower plasma lactate level in healthy controls as compared to lung cancer patients, who typically have a lower HSA concentration. To address this issue, TSP was introduced as a strong HSA-binding competitor. Since TSP exhibits a higher affinity for HSA, its addition results in a saturation of the metabolite-binding sites with TSP and the subsequent dissociation of metabolites from HSA, leading to the correct measurement of the total plasma metabolite concentrations. This is also nicely shown in the study presented in part III, in which lactate was found to be decreased in the preoperative metabolite profile (i.e., the metabolite profile of lung cancer), thereby indicating that the measured lactate concentration without the addition of an HSA-binding competitor rather reflects the differences in HSA concentration between the pre- and postoperative groups and that caution should be taken when describing altered metabolic pathways if no HSA-binding competitor was added to the plasma samples.

A second explanation for interstudy variabilities is related to the normalization method. Many studies performed normalization by calculating the total integration area of the full ^{1}H-NMR spectrum and normalizing the integration value of each integration region towards the sample-specific total [20]. However, statistical analysis may be more strongly influenced by concentration differences in highly abundant plasma metabolites as these dominate the total area. Using this normalization method, plasma metabolite concentrations can only be determined relatively. For this reason, the use of an internal standard is recommended for absolute quantification. Derveaux et al. added a known concentration of MA to each sample, as this would allow for normalization towards MA to obtain peak areas independent of the changes in other signals.

A third explanation for the observed variability in the identified key contributing metabolites could be attributed to the differences in the inclusion criteria of the various studies. While the study discussed in part II includes patients with any stage of lung cancer, where a large part was characterized by metastasis (e.g., >25% stage IV patients in the studies presented in parts I and II), the study discussed in part III only focused on early-stage lung cancers. Since advanced disease stages show more metabolic changes, the plasma metabolite concentrations will also differ from those in the early stages. A study by Larkin et al. confirmed this by demonstrating the ability of plasma metabolites to discriminate early stages from metastasized cancer stages [33]. Here, it was shown that patients with metastatic cancer had increased levels of metabolites such as threonine. Since threonine was not considered to be significant in the last study by Derveaux et al., where only early stages of lung cancer were included, observed differences are likely to be caused by advanced lung cancer stages [21]. This also explains why decreased lipid concentrations (sphingomyelin, phosphatidylcholines, and phospholipids) and increased glucose concentrations in lung cancer plasma were found in multiple publications, including the study discussed in part II, which was absent in part III [18,19,25]. These examples, therefore, further highlight the importance of careful interpretation of the changes in the biochemical pathways, which can heavily depend on the cancer stages of the included patients.

A fourth explanation for different findings can be found in patient-to-patient variations. In most metabolomics studies, a group of (lung) cancer patients is compared with a group of healthy controls or with patients diagnosed with a different type of cancer. While this type of study-set up provides a strong confirmation of the feasibility of differentiating two groups, which is needed for applying the metabolite profile in a diagnostic setting, individual follow-up of patients after treatment remains very challenging. For this reason, the study of Derveaux et al., described in part III, compares pre- and postoperative plasma metabolite profiles of early-stage NSCLC patients via a longitudinal study design (of the

same patients). This way, each patient can serve as their own control, thereby greatly reducing patient-to-patient variabilities. The effect of this longitudinal sample collection was first confirmed by comparing the pre-and postoperative metabolite profiles (i.e., the metabolite signatures of lung cancer and healthy condition, respectively) via OPLS-DA analysis, where the paired character of the samples was not taken into consideration yet. The success of a further reduction in patient-to-patient variation was confirmed using OPLS-EP, in which the paired character of the samples was taken into account.

The importance of all the methodology improvements discussed above is, for example, demonstrated when analyzing the lactate plasma levels in lung cancer. While the lactate concentration in the plasma of lung cancer patients is shown to be decreased in the most recent study by Derveaux et al., contradictory results are found throughout the literature. Zhang et al. compared the metabolite profile of serum from lung cancer patients with healthy controls using a non-targeted ^{1}H-NMR approach and a targeted rapid resolution liquid chromatography [32]. Here, lactate was identified as a potential diagnostic biomarker, but with an increase in concentration in lung cancer. Rocha et al. also found an increased lactate concentration in the plasma of lung cancer patients [25]. The first explanation for these contradictory results is the absence of an HSA-binding competitor. Derveaux et al. demonstrated the high affinity of lactate for HSA, resulting in the strong binding of lactate to HSA [20]. Since lung cancer patients are known to have decreased HSA concentrations, less lactate is able to bind, resulting in an apparent higher concentration in the plasma or serum of lung cancer patients compared to healthy controls with more HSA, and therefore less unbound lactate [34]. The results of these studies, therefore, reflect the decrease in plasma HSA concentration of lung cancer patients, rather than the actual total plasma lactate concentration. An additional reason for these contradictory results can be found in the different disease stages of the included lung cancer patients. While the most recent study by Derveaux et al. only included early-stage NSCLC, some studies included all stages of lung cancer [21]. Puchades-Carrasco et al. compared metabolic biomarkers between early-stage and advanced-stage NSCLC, in which lactate was found to be significantly higher in advanced stages compared to early disease stages, showing the importance of the disease stage in the metabolite profile [24]. Besides disease stage, disease type should also be considered. Zhang et al. included lung cancer patients with either adenosquamous-, adeno-, or small-cell carcinomas [32]. Kowalczyk et al. studied whether the tissue metabolite profile could be used to discriminate between different types of lung cancer and were able to distinguish between adeno- and small-cell carcinoma [35]. A distinction could also be made using blood plasma, although other metabolites seemed to contribute to the observed differences in metabolite profiles. Even though lactate was not identified as a significant contributor in tissue, it is unknown whether plasma analysis would identify lactate as an important metabolite in this study [35]. Therefore, caution should be taken regarding the inclusion of patients with different types of lung cancer if the goal is to unravel the adapted metabolic pathways for a certain type, rather than cancer diagnosis.

Altogether, the study discussed in part III combined the three aforementioned advantages: (i) the addition of TSP to the samples as an HSA-binding competitor, (ii) the inclusion of only early-stage NSCLC patients, and (iii) the reduction in patient-to-patient variability by comparing paired pre- with postoperative plasma metabolite profiles. The study clearly revealed four key indicators for the onset of early-stage NSCLC: lactate, asparagine, acetate, and cysteine.

6. Metabolic Pathways Involved in (Lung) Cancer

As optimized preanalytical conditions were used in the study of part III, resulting in accurate measurements of total plasma metabolite levels, this section focuses on the changes in biochemical pathways due to in-/decreased plasma metabolite levels that were revealed as key contributors in the changed lung cancer metabolite profile.

6.1. Lactate

Cancer cells are characterized by high energy demands due to high proliferation rates, and one of their hallmarks is their ability to reprogram the metabolism, including the glucose and lactate metabolism [36]. Indeed, the study described in part III clearly reveals that lactate levels are decreased in lung cancer plasma. This important finding can be explained by a detailed description of the lung cancer's metabolism. In typical glycolytic cancer cells, this high energy demand is met by increased glucose uptake via the upregulated levels of glucose transporters (GLUTs), especially GLUT3 and GLUT5 in lung cancer cells (Figure 8) [37]. Unlike in normal cells, where glucose is broken down to two molecules of pyruvate which can enter the TCA cycle to produce at least 32 ATP molecules, cancer cells metabolize glucose in a different way. Cancer cells partially metabolize glucose via the nonoxidative catabolism, even in the presence of sufficient amounts of oxygen. Pyruvate is converted into lactate, preventing its use as an intermediate to form acetyl-CoA to drive the TCA cycle, thereby resulting in the production of only 2 ATP molecules per glucose molecule [38]. This phenomenon is also known as the 'Warburg effect' or aerobic fermentation. Even though it was believed that glycolytic cancer cells only used aerobic fermentation, it is known by now that ATP is also produced via oxidative phosphorylation (OXPHOS). Since OXPHOS yields higher levels of ATP production compared to aerobic fermentation (aerobic glycolysis), it is believed that the Warburg effect occurs to support biosynthetic requirements of uncontrolled proliferation, while the OXPHOS produces sufficient energy to support cell growth. This hypothesis is further supported by the fact that a different form of pyruvate kinase (PK) is expressed in lung cancer cells [38]. PK catalyzes the conversion of phosphoenolpyruvate (PEP) into pyruvate, thereby releasing one ATP molecule. In normal cells, the tetrameric form of isoenzyme PKM2 is expressed, which is characterized by a high affinity for PEP. Lung cancer cells, however, express the dimeric form of PKM2, which has a much lower affinity for PEP (Figure 8). Therefore, all glycolytic intermediates preceding PKM2 activity will accumulate and be used for biosynthetic processes to stimulate tumor proliferation such as nucleotide synthesis [38] (Figure 8). This does not imply that pyruvate will no longer be formed at all, but rather to a lesser extent. Pyruvate will then either be converted into lactate or enter the OXPHOS as discussed earlier. The upregulated conversion to lactate in glycolytic cancer cells further leads to increased amounts of intracellular lactate production and excretion via monocarboxylate transporter 4 (MCT4) (Figure 8). This increased excretion of lactate into the microenvironment results in acidification, immune cell function modulation, extracellular drug deactivation, and promotion of invasion and metastasis [36,38]. Lactate, produced by the glycolytic cancer cells, can, however, be taken up from the microenvironment by oxidative cancer cells via MCT1 transporters in order to be used as a substrate for increased gluconeogenesis since it is the primary source of a gluconeogenic precursor [39,40]. This transport of lactate between glycolytic and oxidative cancer cells is also referred to as the "lactate shuttle" [41]. Studies have shown that one of the key enzymes involved in gluconeogenesis, phosphoenolpyruvate carboxykinase (PEPCK), is responsible for the replenishment of glycolytic pools in cancer cells, thereby supporting cancer cell growth. Leithner et al. showed that one of the isoforms of this enzyme, PCK2 (or PEPCK2), was shown to be actively expressed by lung cancer cells and that expression levels were enhanced under low-glucose conditions [39]. This principle was confirmed since PEPCK inhibition led to a reduction in tumor growth [42]. The gluconeogenesis is even further supported in lung cancer by the upregulation of pyruvate carboxylase (PC) since this enzyme is responsible for the transformation of pyruvate to oxaloacetate [43]. Oxaloacetate will eventually be converted to PEP by the increased levels of PCK2. In this way, oxidative lung cancer cells can convert lactate into PEP via the upregulation of PC and PCK2 (Figure 8). PEP can eventually donate its carbons to the gluconeogenesis pathway, thereby providing energy in a glucose-independent manner. This process, therefore, reflects the metabolic symbiosis between glycolytic and oxidative tumor cells in order to survive and stimulate cancer growth.
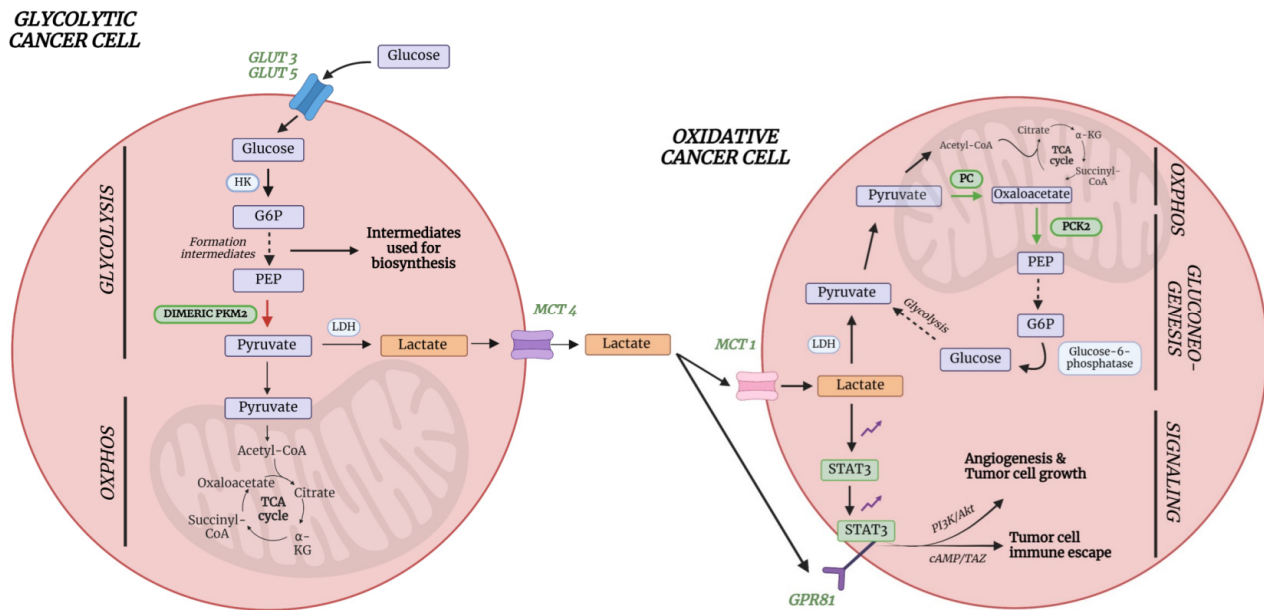
Figure 8. Metabolic symbiosis between glycolytic and oxidative lung cancer cells showing the lactate shuttle. Glycolytic cancer cells use aerobic glycolysis (aerobic fermentation) to obtain intermediates for biosynthesis, while oxidative phosphorylation is used to yield energy for further tumor growth. Nevertheless, glycolysis is mainly increased in the glycolytic cancer cells, while oxidative phosphorylation is decreased. As aerobic glycolysis is used, lactate is formed and released in the microenvironment by MCT4 transporters. Lactate can be taken up by oxidative cancer cells ("Lactate shuttle") by the MCT1 transporters for gluconeogenesis. The glucose that is eventually formed will undergo glycolysis to stimulate biosynthesis via its intermediates and partially enter the TCA cycle after conversion to pyruvate to yield energy. Additionally, lactate will bind to the GRP81 receptor, which is upregulated by lactate itself via STAT3 in lung cancer cells, resulting in angiogenesis, tumor cell growth, and tumor cell immune escape. Enzymes, receptor labels, and arrows in green represent an upregulation, while red represents a downregulation. Akt, protein kinase B; α-KG, alpha-ketoglutarate; cAMP, cyclic adenosine monophosphate; GLUT3, glucose transporter 3; GLUT5, glucose transporter 5; GPR81, G-protein coupled receptor 81; G6P, glucose-6-phosphate; HK, hexokinase; LDH; lactate dehydrogenase; MCT1, monocarboxylate transporter 1; MCT4, monocarboxylate transporter 4; OXPHOS, oxidative phosphorylation; PC, pyruvate carboxylase; PEP, phosphoenolpyruvate; PCK2, phosphoenolpyruvate carboxykinase 2; PKM2, pyruvate kinase M2; PI3K, phosphoinositide 3-kinase; STAT3, signal transducer and activator of transcription 3; TAZ, transcriptional coactivator with PDZ-binding motif; TCA, tricarboxylic acid cycle. Figure created with BioRender.com.

Besides the fact that lactate can be used as an energy source and precursor for gluconeogenesis, it can also be used as a signaling molecule for tumor survival and progression. Lactate is able to bind to the membrane-bound receptor G protein-coupled receptor 81 (GPR81), which is typically upregulated in cancer cells [44] (Figure 8). This upregulation in lung cancer cells is caused by lactate via the signal transducer and activator of transcription 3 (STAT3). Lactate upregulates the transcriptional factor Snail, resulting in the assembly of the Snail/EZH2/STAT3 complex, which enhances STAT3 activity. STAT3 will in turn bind to the GPR81 promotor, activating its expression in cancer cells [45]. The binding of lactate to this receptor could eventually lead to the activation of the phosphatidylinositol 3-kinase (Pi3K)/protein kinase B (Akt) pathway, resulting in the release of amphiregulin, which stimulates angiogenesis of tumor endothelial cells, tumor cell growth, and proliferation (Figure 8) [44]. Another consequence of GPR81 binding in lung cancer is the induction of the programmed cell death protein 1 ligand (PD-L1), causing tumor cell immune escape [46] (Figure 8). Therefore, despite increased production and excretion rates, a decrease in plasma lactate concentrations can be observed.

6.2. Acetate and Amino Acids Cysteine and Asparagine

The most recent paper by Derveaux et al. showed that lactate, cysteine, acetate, and asparagine were the most significant metabolites in the differentiation between early-stage lung cancer and healthy conditions [21]. Acetate was found to be increased in the plasma of lung cancer patients, while cysteine and asparagine were shown to be decreased. This will be discussed more in detail since the study took the following into account: (i) limited patient-to-patient variability (paired samples); (ii) use of a new methodology that more accurately represents the total metabolite concentrations via the addition of TSP as a binding competitor and MA as an internal standard; and (iii) only early stages of NSCLC (I-IIIa) were included (instead of all NSCLC stages), thereby limiting variability.

Acetate metabolism is an important pathway in many cancers, including lung cancer, since they are characterized by an increased expression of acetate-metabolizing enzymes like acetyl-CoA synthetases [47]. Lung cancer is characterized by an increased expression of acetyl-CoA synthetase 2 (ACSS2), which catalyzes the ATP-dependent ligation of acetate and CoA to produce acetyl-CoA. Acetyl-CoA will then be used for energy production, but also as a substrate for other biochemical reactions such as the synthesis of ketones and sterols [47]. This production of energy substrates enables further stimulation of tumor growth. This was further confirmed by Comerford et al., who showed that ACSS2 knockdown in human patient tumor cell lines led to growth inhibition [48]. Another important enzyme in lung cancer is ATP-citrate lyase (ACLY), a key enzyme of de novo fatty acid synthesis which is responsible for the transformation of TCA-derived citrate into acetyl-CoA [49]. Migita et al. showed an increase in ACLY expression in human lung adenocarcinoma cells, indicating an increased formation of acetyl-CoA, which can be used as a precursor for fatty acid synthesis, cholesterol synthesis, and energy production. The importance of this enzyme was demonstrated by inhibition via RNA interference, which induced tumoral growth arrest. Even though an increased acetate consumption is seen in cancer, an increase in plasma concentration can be observed in lung cancer patients. Research has demonstrated a pathway involved in de novo acetate production from pyruvate, which was pronounced at moments of overflow metabolism [50]. Overflow metabolism is defined as a situation in which the metabolic supply outpaces the metabolic demand, resulting in an accumulation of intermediate products [50]. This is typically observed in cancers, where an incomplete metabolism occurs together with increased glucose uptake by cancer cells, yielding metabolic intermediates, such as acetate, into the extracellular space. This excessive amount of acetate that is released by nutrient-rich tumor cells can eventually be consumed by nutrient-poor tumor cells, allowing them to function and grow in poor vascular environmental conditions. Acetate, therefore, plays a key molecule in the symbiotic interaction that occurs between nutrient-poor and -rich tumor cells [50]. Another explanation for the increased acetate plasma concentration can be found within systemic metabolism. The body itself can also act on cancer development, thereby contributing to altered plasma metabolite concentrations. During cancer, increased levels of acyl-CoA thioesterase 12 (ACOT12), an enzyme which generates acetate from acetyl-CoA, can be found in the liver. This process is activated in conditions of metabolic stress, resulting in the formation and excretion of acetate in the extracellular space [51].

Also, the cysteine plasma concentration shows a significant alteration in lung cancer compared to normal plasma. Contrary to the increased acetate levels, decreased cysteine plasma levels are found in lung cancer. Cysteine is of great importance in energy production, cellular biosynthesis, enzyme catalysis, and redox metabolism [52]. During tumor development, there is an increased demand for cysteine to elevate glutathione production to protect cancer cells from death due to increased, accumulated oxidative stress levels. Cysteine is of great importance, together with glycine and glutamate, for glutathione formation by glutamyl-cysteine ligase (GCL) [52]. Cysteine is transported into the cell using both the excitatory amino acid transporter 3 (EAAT3/SCL1A1) and the alanine–serine–cysteine transporter 1 and 2 (ASCT1/2) [53]. In lung cancer, increased cysteine uptake is enabled by the overexpression of ASCT2 and EAAT3 [54]. Besides its role in ROS protection, cysteine is also used as a carbon source during energy production by contributing to the TCA cycle

after its conversion to pyruvate or α-ketoglutarate [55]. Even though many cancer types show an increase in de novo synthesis of cysteine, lung cancer cells show an inactivated pathway during tumorigenesis [52]. Despite the inactivation of cysteine synthesis, cysteine uptake and metabolism remain. Taken together, a decrease in cysteine plasma concentration is observed in lung cancer due to the increased metabolic needs combined with decreased de novo synthesis.

Finally, asparagine was shown to be decreased in the plasma of lung cancer patients. The asparagine metabolism is closely related to the glutamine metabolism. Glutamine is one of the major precursors for de novo biosynthesis of nucleotides as well as other nonessential amino acids [56]. Once glutamine donates its amino group for nucleotide synthesis, in turns into glutamate, which eventually can be used for the synthesis of nonessential amino acid derivatives such as α-ketoglutarate. In this way, α-ketoglutarate can enter the TCA cycle, yielding additional energy for further cancer cell proliferation and survival. Within glutamine-deprived cells, which are frequently observed in cancer cells under physiological conditions, asparagine plays an important role in maintaining cancer cell viability. Just like glutamine, asparagine functions as a precursor for the TCA cycle. Asparagine can be converted into aspartate by asparaginase, which can be converted by aminotransferases into oxaloacetate, which can enter the TCA cycle, providing new energy for cancer cell proliferation. Therefore, asparagine was originally thought to prevent glutamine depletion-induced apoptosis by acting as a precursor substrate for anaplerosis (replenishment of TCA cycle intermediates) [56]. Nevertheless, this hypothesis was challenged as it was shown that inhibition of citrate synthase (CS), the first enzyme of the TCA cycle, prevented glutamine depletion-induced apoptosis. Additionally, asparagine addition did not result in anaplerosis, indicating that the survival-stimulating effect of asparagine was not caused by acting as a precursor for the TCA cycle in cancer cells. Asparagine is, therefore, not catabolized to act as a carbon donor, but it is taken up directly from the extracellular environment to protect cancer cells from glutamine depletion-induced apoptosis and to maintain protein synthesis [57]. Zhang et al. showed that glutamine-deprived cancer cells first undergo cell-cycle arrest, but that translation of the stress-response RNAs, including C/EBP-homologous protein (CHOP), results in the actual activation of cellular apoptosis [56]. To prevent this glutamine depletion-induced apoptosis, asparagine steps in as it is able to suppress the activation of CHOP. The critical role of asparagine was confirmed through the inhibition of asparagine synthetase (ASNS), which resulted in the induction of apoptosis. In addition, Jiang et al. showed that exogenous asparagine is able to induce glutamine synthetase (GLUL) accumulation at a post-transcriptional level, enabling de novo glutamine synthesis [58].

7. Conclusions and Future Perspectives

Early-stage plasma biomarkers have great potential to enable early detection of lung cancer, thereby providing a faster diagnosis and improved survival rates. The above-mentioned studies confirmed changes in the plasma metabolite profile upon lung cancer development, enabling to distinguish these patients from the controls. Furthermore, the differentiating OPLS-DA and OPLS-EP models in these studies identified several metabolites as key contributors (i.e., lactate, cysteine, asparagine, and acetate) and therefore as potential plasma biomarkers for lung cancer.

^{1}H-NMR is currently used as the golden standard for metabolomics research since it is characterized by multiple advantages such as its simple sample preparation without the need for extractions, noninvasive character, high reproducibility (>98%), short run time, automatization possibilities and only minor matrix effects. The main challenges that remtain include the need for (international) standardized protocols to decrease variabilities caused by alterations in working procedures and to enable a more detailed comparison of study results. An essential step, therefore, is to standardize the incorporation of (i) TSP as an HSA-binding competitor and (ii) a suitable internal standard such as MA. As TSP enables the dissociation of HSA-bound metabolites, representing the actual plasma metabolite concentrations, and an internal standard enables normalization independent from changes in metabolite levels, both practices are required for accurate and quantitative metabolic

analysis. Besides the need for standardized protocols, it is crucial to sufficiently use large sample sizes to prevent overfitting when using multivariate statistics, given the large number of variables in metabolomics studies.

These metabolomic studies led to several advantages, i.e., the ability to diagnose lung cancer in a non-invasive, efficient manner, and potential future applications of the plasma metabolite changes were shown throughout the different studies discussed. Louis et al. were able to show the potential of plasma metabolites as biomarkers for lung cancer diagnosis and screening since clear differences were observed between plasma metabolite profiles of lung cancer patients compared to controls [18]. Additionally, the discriminating power of plasma metabolites was shown by their ability to distinguish lung cancer from breast cancer based on different metabolic alterations [19,29]. The last study by Derveaux et al. showed a clear "metabolic shift" in plasma metabolite concentrations between pre- and postoperative states [21]. In general, we aim to use ^{1}H-NMR to provide support in three stages of clinical decision making: (i) screening via the combination of ^{1}H-NMR and LDCT, (ii) ^{1}H-NMR to complement PET, and (iii) the use of ^{1}H-NMR for disease prediction and monitoring. Regarding the first aim, we would like to combine ^{1}H-NMR and LDCT for screening a large high-risk population. Secondly, we would like to use ^{1}H-NMR to complement the PET scan. As PET scans often fail to differentiate between malignant and non-malignant PET-positive lesions, a high number of false positives is observed. Combining PET scans with ^{1}H-NMR could overcome this by taking the plasma metabolomics into account. Thirdly, we would ultimately also like to use ^{1}H-NMR as a tool for disease prediction and monitoring. Since the plasma metabolite profile changes after disease recovery, plasma biomarkers are a promising, complementary tool to evaluate therapy response, follow-up, and disease recurrence. Plasma biomarkers could even be used in the future to identify disruptions in metabolic pathways with greater detail, allowing for targeted therapeutic interventions as knowledge of altered metabolic processes continues to expand. In summary, it becomes increasingly clear that NMR metabolomics might be able to pave the way towards providing support in clinical decision making regarding early cancer diagnosis via screening (of encumbered populations), complementing PET, assessing therapy response and follow-up monitoring, and potentially predicting disease recurrence.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A
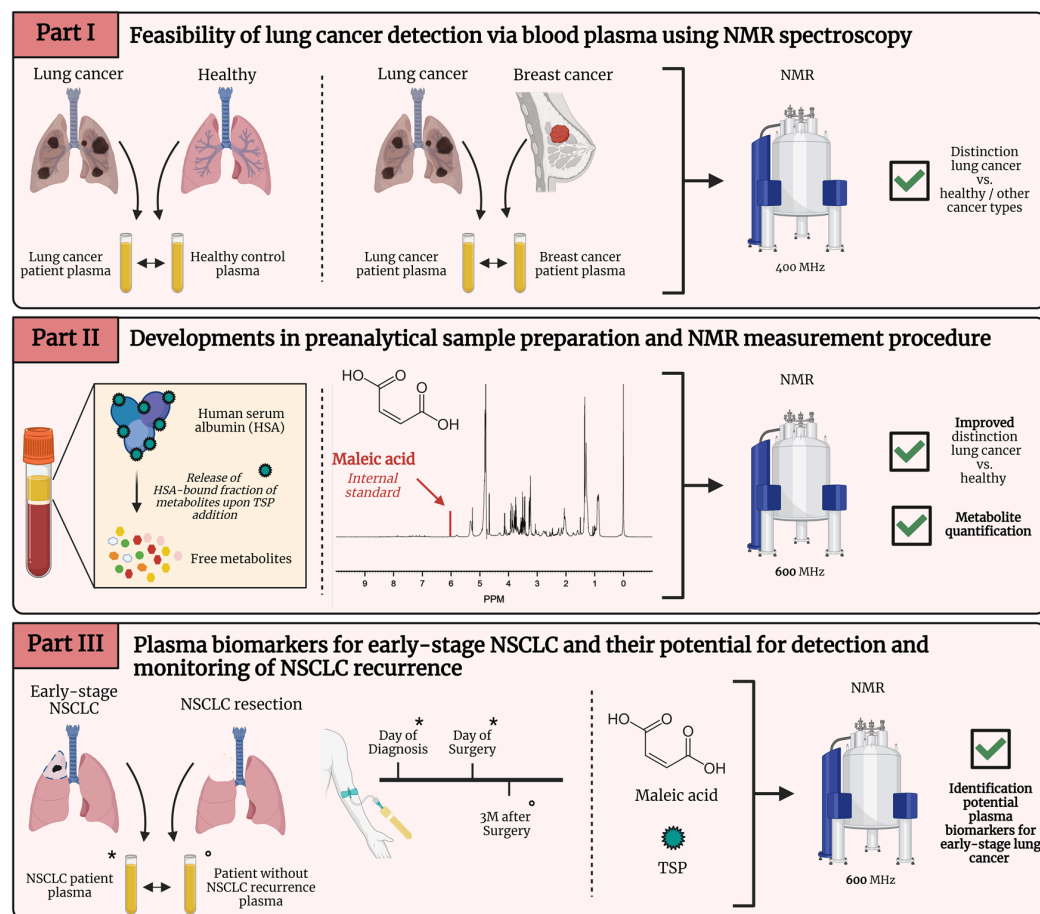
Figure A1. Graphical overview of parts I, II, and III. * indicate preoperative plasma samples, ° indicate postoperative plasma samples. MHz, megahertz; NMR, nuclear magnetic resonance; NSCLC, non-small cell lung cancer; TSP, trimethylsilyl-2,2,3,3-tetradeuteropropionic acid. Figure created with BioRender.com.

Appendix B

For the fundamentals of ^{1}H-NMR, the interested reader is referred to the book by David G. Gadian titled *NMR and Its Application to Living Systems* [59].

Appendix B.1. Sample Collection

Blood samples were collected in lithium heparin tubes and stored at 4 °C for 10 min. Plasma was isolated after centrifugation at $1600\times g$ for 15 min (swinging bucket centrifuge), aliquoted, and stored at −80 °C until NMR analysis.

Appendix B.2. Sample Preparation

Stock solutions of 1 M K_2HPO_4 and 1 M KH_2PO_4 were prepared in D_2O and further diluted to a final concentration of 0.15 M potassium phosphate and calibrated to a pH of 7.4. Finally, 8 mM TSP (HSA-binding competitor) and 107.70 µM MA (internal standard) were added to the 0.15 M phosphate buffer. Afterward, plasma aliquots were centrifuged at $13,000\times g$ for 4 min at 4 °C, and 350 µL of plasma was added to 350 µL of 0.15 M potassium phosphate buffer at pH 7.4 in D_2O. This resulted in a 700 µL sample containing 4 mM TSP and 53.85 µM MA.

Appendix B.3. Metabolite Spiking

To assign all peaks of the plasma ^{1}H-NMR spectrum, reference plasma samples (obtained from a healthy person) were spiked with a specific metabolite and analyzed with ^{1}H-NMR. This was repeated for every metabolite separately. If a specific spiked metabolite is present in the sample, its corresponding signal(s) will show an increase in the NMR spectrum. This enables the identification of the metabolite. In the studies of Louis et al., a 400 MHz NMR spectrometer was used, and based on the spiking information, the ^{1}H-NMR spectrum was divided into 110 integration regions. In the study by Derveaux et al., a 600 MHz NMR spectrometer was used, and based on the metabolite spiking information, the ^{1}H-NMR spectrum was divided into 237 integration regions, thereby assigning all the signals of 62 plasma metabolites. An example is shown in Figure A2.

Appendix B.4. NMR Data Acquisition

After thermal sample stabilization for 5 min at 25 °C, ^{1}H-NMR spectra were recorded at 25 °C with 96 scans. Spectra were obtained using a Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence, which suppresses broad signals of macromolecules in the ^{1}H-NMR spectrum.

Appendix B.5. NMR Spectra Processing

All spectra were phased manually and baseline-corrected. In the study by Derveaux et al., the 237 integration regions were normalized to the peak area of the maleic acid internal standard. The integration regions are the variables for the multivariate statistics, and the value of each variable is determined by the area under the peak of the integration region (and, therefore, directly correlated with the concentration of a specific metabolite).
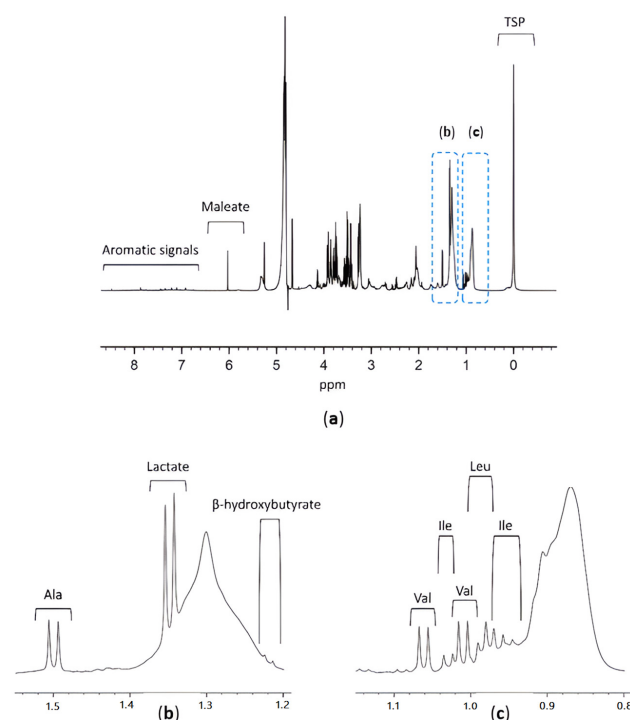
Figure A2. (**a**) Example of the ^{1}H-NMR spectrum of a healthy control using a 600 MHz NMR spectrometer, with TSP as HSA-binding competitor and maleate as the internal standard for normalization. Chemical shift values are reported in parts per million (ppm). The area under the peak corresponds with the molar amount of hydrogen atoms and, therefore, the metabolite concentration. (**b**) Magnification of the 1.2–1.5 ppm region, where alanine, lactate, and β-hydroxybutyrate are identified using metabolite spiking. (**c**) Magnification of the 0.8–1.1 ppm region, where valine, isoleucine, and leucine are identified using metabolite spiking. Ala, alanine; Ile, isoleucine; Leu, leucine; TSP, trimethylsilyl-2,2,3,3-tetradeuteropropionic acid; Val, valine. Figure adapted from reference [15].

References

1. Hirsch, F.R.; Scagliotti, G.V.; Mulshine, J.L.; Kwon, R.; Curran, W.J.; Wu, Y.-L.; Paz-Ares, L. Lung cancer: Current therapies and new targeted treatments. *Lancet* **2017**, *389*, 299–311. [CrossRef] [PubMed]
2. Wood, D.E.; Kazerooni, E.A.; Baum, S.L.; Eapen, G.A.; Ettinger, D.S.; Hou, L.; Jackman, D.M.; Klippenstein, D.; Kumar, R.; Lackner, R.P.; et al. Lung Cancer Screening, Version 3.2018, NCCN Clinical Practice Guidelines in Oncology. *J. Natl. Compr. Cancer Netw.* **2018**, *16*, 412–441. [CrossRef] [PubMed]
3. Lancaster, H.; Heuvelmans, M.A.; Oudkerk, M. Low-dose computed tomography lung cancer screening: Clinical evidence and implementation research. *J. Intern. Med.* **2022**, *292*, 68–80. [CrossRef] [PubMed]
4. The National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N. Engl. J. Med.* **2011**, *365*, 395–409. [CrossRef] [PubMed]
5. Duffy, S.W.; Field, J. Mortality Reduction with Low-Dose CT Screening for Lung Cancer. *N. Engl. J. Med.* **2020**, *382*, 572–573. [CrossRef] [PubMed]
6. Oudkerk, M.; Liu, S.; Heuvelmans, M.A.; Walter, J.E.; Field, J.K. Lung cancer LDCT screening and mortality reduction—Evidence, pitfalls and future perspectives. *Nat. Rev. Clin. Oncol.* **2021**, *18*, 135–151. [CrossRef] [PubMed]
7. Berg, C.D.; Aberle, D.R.; Wood, D.E. Lung Cancer Screening: Promise and Pitfalls. *Am. Soc. Clin. Oncol. Educ. Book* **2012**, *32*, 450–457. [CrossRef] [PubMed]
8. Farwell, M.D.; Pryma Da Fau-Mankoff, D.A.; Mankoff, D.A. PET/CT imaging in cancer: Current applications and future directions. *Cancer* **2014**, *120*, 3433–3445. [CrossRef] [PubMed]
9. Deppen, S.A.; Blume, J.D.; Kensinger, C.D.; Morgan, A.M.; Aldrich, M.C.; Massion, P.P.; Walker, R.C.; McPheeters, M.L.; Putnam, J.B., Jr.; Grogan, E.L. Accuracy of FDG-PET to diagnose lung cancer in areas with infectious lung disease: A meta-analysis. *JAMA* **2014**, *312*, 1227–1236. [CrossRef]
10. Rosenbaum, S.J.; Lind, T.; Antoch, G.; Bockisch, A. False-Positive FDG PET Uptake−the Role of PET/CT. *Eur. Radiol.* **2006**, *16*, 1054–1065. [CrossRef]
11. Steven, B.Z.; Elizabeth, T.L.; Christopher, G.S.; David, H.A.; Paul, L.H.; Gregory, J.K.; Larry, G.K.; Leah, M.B. Preoperative PET and the Reduction of Unnecessary Surgery Among Newly Diagnosed Lung Cancer Patients in a Community Setting. *J. Nucl. Med.* **2014**, *55*, 379. [CrossRef]
12. Vanhove, K.; Giesen, P.; Owokotomo, O.E.; Mesotten, L.; Louis, E.; Shkedy, Z.; Thomeer, M.; Adriaensens, P. The plasma glutamate concentration as a complementary tool to differentiate benign PET-positive lung lesions from lung cancer. *BMC Cancer* **2018**, *18*, 868. [CrossRef] [PubMed]
13. Seijo, L.M.; Peled, N.; Ajona, D.; Boeri, M.; Field, J.K.; Sozzi, G.; Pio, R.; Zulueta, J.J.; Spira, A.; Massion, P.P.; et al. Biomarkers in Lung Cancer Screening: Achievements, Promises, and Challenges. *J. Thorac. Oncol.* **2019**, *14*, 343–357. [CrossRef] [PubMed]
14. Clish, C.B. Metabolomics: An emerging but powerful tool for precision medicine. *Cold Spring Harb. Mol. Case Stud.* **2015**, *1*, a000588. [CrossRef]
15. Vanhove, K.; Derveaux, E.; Mesotten, L.; Thomeer, M.; Criel, M.; Mariën, H.; Adriaensens, P. Unraveling the Rewired Metabolism in Lung Cancer Using Quantitative NMR Metabolomics. *Int. J. Mol. Sci.* **2022**, *23*, 5602. [CrossRef]
16. Noreldeen, H.A.A.; Liu, X.; Xu, G. Metabolomics of lung cancer: Analytical platforms and their applications. *J. Sep. Sci.* **2020**, *43*, 120–133. [CrossRef]
17. Pascale, R.; Calvisi, D.F.; Simile, M.M.; Feo, C.; Feo, F. The Warburg Effect 97 Years after Its Discovery. *Cancers* **2020**, *12*, 2819. [CrossRef] [PubMed]
18. Louis, E.; Adriaensens, P.; Guedens, W.; Bigirumurame, T.; Baeten, K.; Vanhove, K.; Vandeurzen, K.; Darquennes, K.; Vansteenkiste, J.; Dooms, C.; et al. Detection of Lung Cancer through Metabolic Changes Measured in Blood Plasma. *J. Thorac. Oncol.* **2016**, *11*, 516–523. [CrossRef]
19. Louis, E.; Adriaensens, P.; Guedens, W.; Vanhove, K.; Vandeurzen, K.; Darquennes, K.; Vansteenkiste, J.; Dooms, C.; de Jonge, E.; Thomeer, M.; et al. Metabolic phenotyping of human blood plasma: A powerful tool to discriminate between cancer types? *Ann. Oncol.* **2016**, *27*, 178–184. [CrossRef]
20. Derveaux, E.; Thomeer, M.; Mesotten, L.; Reekmans, G.; Adriaensens, P. Detection of Lung Cancer via Blood Plasma and (1)H-NMR Metabolomics: Validation by a Semi-Targeted and Quantitative Approach Using a Protein-Binding Competitor. *Metabolites* **2021**, *11*, 537. [CrossRef]
21. Derveaux, E.; Geubbelmans, M.; Criel, M.; Demedts, I.; Himpe, U.; Tournoy, K.; Vercauter, P.; Johansson, E.; Valkenborg, D.; Vanhove, K.; et al. NMR-Metabolomics Reveals a Metabolic Shift after Surgical Resection of Non-Small Cell Lung Cancer. *Cancers* **2023**, *15*, 2127. [CrossRef] [PubMed]
22. Worley, B.; Powers, R. Multivariate Analysis in Metabolomics. *Curr. Metabolomics* **2013**, *1*, 92–107. [CrossRef] [PubMed]
23. Eriksson, L.; Byrne, T.; Johansson, E.; Trygg, J.; Vikström, C. *Multi- and Megavariate Data Analysis Basic Principles and Applications*; Umetrics Academy: Umeå, Sweden, 2013.
24. 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]

25. Rocha, C.M.; Carrola, J.; Barros, A.S.; Gil, A.M.; Goodfellow, B.J.; Carreira, I.M.; Bernardo, J.; Gomes, A.; Sousa, V.; Carvalho, L.; et al. Metabolic signatures of lung cancer in biofluids: NMR-based metabonomics of blood plasma. *J. Proteome Res.* **2011**, *10*, 4314–4324. [CrossRef] [PubMed]

26. Xu, Y.; Goodacre, R. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. *J. Anal. Test.* **2018**, *2*, 249–262. [CrossRef] [PubMed]

27. Guo, L.; Liu, J.; Lu, R. Subsampling bias and the best-discrepancy systematic cross validation. *Sci. China Math.* **2021**, *64*, 197–210. [CrossRef]

28. Carrola, J.; Rocha, C.M.; Barros, A.S.; Gil, A.M.; Goodfellow, B.J.; Carreira, I.M.; Bernardo, J.; Gomes, A.; Sousa, V.; Carvalho, L.; et al. Metabolic signatures of lung cancer in biofluids: NMR-based metabonomics of urine. *J. Proteome Res.* **2011**, *10*, 221–230. [CrossRef] [PubMed]

29. Louis, R.; Louis, E.; Stinkens, K.; Mesotten, L.; de Jonge, E.; Thomeer, M.; Caenepeel, P.; Adriaensens, P. Metabolic Phenotyping of Blood Plasma by Proton Nuclear Magnetic Resonance to Discriminate between Colorectal Cancer, Breast Cancer and Lung cancer. *Metabolomics* **2016**, *6*, 187. [CrossRef]

30. Pearce, J.T.M.; Athersuch, T.J.; Ebbels, T.M.D.; Lindon, J.C.; Nicholson, J.K.; Keun, H.C. Robust algorithms for automated chemical shift calibration of 1D 1H NMR spectra of blood serum. *Anal. Chem.* **2008**, *80*, 7158–7162. [CrossRef]

31. Jonsson, P.; Wuolikainen, A.; Thysell, E.; Chorell, E.; Stattin, P.; Wikström, P.; Antti, H. Constrained randomization and multivariate effect projections improve information extraction and biomarker pattern discovery in metabolomics studies involving dependent samples. *Metabolomics* **2015**, *11*, 1667–1678. [CrossRef]

32. Zhang, X.; Zhu, X.; Wang, C.; Zhang, H.; Cai, Z. Non-targeted and targeted metabolomics approaches to diagnosing lung cancer and predicting patient prognosis. *Oncotarget* **2016**, *7*, 63437–63448. [CrossRef] [PubMed]

33. Larkin, J.; Anthony, S.; Johanssen, V.; Yeo, T.; Sealey, M.; Yates, A.G.; Smith, C.; Claridge, T.; Nicholson, B.D.; Moreland, J.; et al. Metabolomic Biomarkers in Blood Samples Identify Cancers in a Mixed Population of Patients with Nonspecific Symptoms. *Clin. Cancer Res.* **2022**, *28*, 1651–1661. [CrossRef] [PubMed]

34. Gupta, D.; Lis, C.G. Pretreatment serum albumin as a predictor of cancer survival: A systematic review of the epidemiological literature. *Nutr. J.* **2010**, *9*, 69. [CrossRef] [PubMed]

35. Kowalczyk, T.; Kisluk, J.; Pietrowska, K.; Godzien, J.; Kozlowski, M.; Reszeć, J.; Sierko, E.; Naumnik, W.; Mróz, R.; Moniuszko, M.; et al. The Ability of Metabolomics to Discriminate Non-Small-Cell Lung Cancer Subtypes Depends on the Stage of the Disease and the Type of Material Studied. *Cancers* **2021**, *13*, 3314. [CrossRef] [PubMed]

36. Ramteke, P.; Deb, A.; Shepal, V.; Bhat, M.K. Hyperglycemia Associated Metabolic and Molecular Alterations in Cancer Risk, Progression, Treatment, and Mortality. *Cancers* **2019**, *11*, 1402. [CrossRef] [PubMed]

37. Kannampuzha, S.; Mukherjee, A.; Wanjari, U.; Gopalakrishnan, A.; Murali, R.; Namachivayam, A.; Renu, K.; Dey, A.; Vellingiri, B.; Madhyastha, H.; et al. A Systematic Role of Metabolomics, Metabolic Pathways, and Chemical Metabolism in Lung Cancer. *Vaccines* **2023**, *11*, 381. [CrossRef] [PubMed]

38. Vanhove, K.; Graulus, G.J.; Mesotten, L.; Thomeer, M.; Derveaux, E.; Noben, J.P.; Guedens, W.; Adriaensens, P. The Metabolic Landscape of Lung Cancer: New Insights in a Disturbed Glucose Metabolism. *Front. Oncol.* **2019**, *9*, 1215. [CrossRef] [PubMed]

39. Leithner, K.; Hrzenjak, A.; Trötzmüller, M.; Moustafa, T.; Köfeler, H.C.; Wohlkoenig, C.; Stacher, E.; Lindenmann, J.; Harris, A.L.; Olschewski, A.; et al. PCK2 activation mediates an adaptive response to glucose depletion in lung cancer. *Oncogene* **2015**, *34*, 1044–1050. [CrossRef] [PubMed]

40. Mendes, C.; Serpa, J. Revisiting lactate dynamics in cancer—A metabolic expertise or an alternative attempt to survive? *J. Mol. Med.* **2020**, *98*, 1397–1414. [CrossRef]

41. Wang, X.; Liu, H.; Ni, Y.; Shen, P.; Han, X. Lactate shuttle: From substance exchange to regulatory mechanism. *Human Cell* **2022**, *35*, 1–14. [CrossRef]

42. Smolle, E.; Leko, P.; Stacher-Priehse, E.; Brcic, L.; El-Heliebi, A.; Hofmann, L.; Quehenberger, F.; Hrzenjak, A.; Popper, H.H.; Olschewski, H.; et al. Distribution and prognostic significance of gluconeogenesis and glycolysis in lung cancer. *Cancers* **2022**, *14*, 5344. [CrossRef] [PubMed]

43. Kiesel, V.A.; Sheeley, M.P.; Coleman, M.F.; Cotul, E.K.; Donkin, S.S.; Hursting, S.D.; Wendt, M.K.; Teegarden, D. Pyruvate carboxylase and cancer progression. *Cancer Metab.* **2021**, *9*, 20. [CrossRef] [PubMed]

44. Li, Z.; Cui, J. Targeting the lactic acid metabolic pathway for antitumor therapy. *Mol. Ther. Oncolytics* **2023**, *31*, 100740. [CrossRef] [PubMed]

45. Xie, Q.; Zhu, Z.; He, Y.; Zhang, Z.; Zhang, Y.; Wang, Y.; Luo, J.; Peng, T.; Cheng, F.; Gao, J.; et al. A lactate-induced Snail/STAT3 pathway drives GPR81 expression in lung cancer cells. *Biochim. Biophys. Acta Mol. Basis Dis.* **2020**, *1866*, 165576. [CrossRef] [PubMed]

46. Feng, J.; Yang, H.; Zhang, Y.; Wei, H.; Zhu, Z.; Zhu, B.; Yang, M.; Cao, W.; Wang, L.; Wu, Z. Tumor cell-derived lactate induces TAZ-dependent upregulation of PD-L1 through GPR81 in human lung cancer cells. *Oncogene* **2017**, *36*, 5829–5839. [CrossRef] [PubMed]

47. Schug, Z.T.; Vande Voorde, J.; Gottlieb, E. The metabolic fate of acetate in cancer. *Nat. Rev. Cancer* **2016**, *16*, 708–717. [CrossRef] [PubMed]

48. Comerford, S.A.; Huang, Z.; Du, X.; Wang, Y.; Cai, L.; Witkiewicz, A.K.; Walters, H.; Tantawy, M.N.; Fu, A.; Manning, H.C.; et al. Acetate dependence of tumors. *Cell* **2014**, *159*, 1591–1602. [CrossRef] [PubMed]

49. Migita, T.; Narita, T.; Nomura, K.; Miyagi, E.; Inazuka, F.; Matsuura, M.; Ushijima, M.; Mashima, T.; Seimiya, H.; Satoh, Y.; et al. ATP citrate lyase: Activation and therapeutic implications in non-small cell lung cancer. *Cancer Res.* **2008**, *68*, 8547–8554. [CrossRef]

50. Bose, S.; Ramesh, V.; Locasale, J.W. Acetate Metabolism in Physiology, Cancer, and Beyond. *Trends Cell Biol.* **2019**, *29*, 695–703. [CrossRef]

51. Horibata, Y.; Ando, H.; Itoh, M.; Sugimoto, H. Enzymatic and transcriptional regulation of the cytoplasmic acetyl-CoA hydrolase ACOT12. *J. Lipid Res.* **2013**, *54*, 2049–2059. [CrossRef]

52. Yoon, S.; Combs, J.; Falzone, A.; Prieto-Farigua, N.; Caldwell, S.; Ackerman, H.; Flores, E.; DeNicola, G. Comprehensive Metabolic Tracing Reveals the Origin and Catabolism of Cysteine in Mammalian Tissues and Tumors. *Cancer Res.* **2023**, *83*, 1426–1442. [CrossRef] [PubMed]

53. Watts, S.D.; Torres-Salazar, D.; Divito, C.B.; Amara, S.G. Cysteine transport through excitatory amino acid transporter 3 (EAAT3). *PLoS ONE* **2014**, *9*, e109245. [CrossRef] [PubMed]

54. García-Gaytán, A.C.; Hernández-Abrego, A.; Díaz-Muñoz, M.; Méndez, I. Glutamatergic system components as potential biomarkers and therapeutic targets in cancer in non-neural organs. *Front. Endocrinol.* **2022**, *13*, 1029210. [CrossRef] [PubMed]

55. Bonifácio, V.D.B.; Pereira, S.A.; Serpa, J.; Vicente, J.B. Cysteine metabolic circuitries: Druggable targets in cancer. *Br. J. Cancer* **2021**, *124*, 862–879. [CrossRef] [PubMed]

56. Zhang, J.; Fan, J.; Venneti, S.; Cross, J.R.; Takagi, T.; Bhinder, B.; Djaballah, H.; Kanai, M.; Cheng, E.H.; Judkins, A.R.; et al. Asparagine plays a critical role in regulating cellular adaptation to glutamine depletion. *Mol. Cell* **2014**, *56*, 205–218. [CrossRef] [PubMed]

57. Pavlova, N.N.; Hui, S.; Ghergurovich, J.M.; Fan, J.; Intlekofer, A.M.; White, R.M.; Rabinowitz, J.D.; Thompson, C.B.; Zhang, J. As Extracellular Glutamine Levels Decline, Asparagine Becomes an Essential Amino Acid. *Cell Metab.* **2018**, *27*, 428.e5–438.e5. [CrossRef]

58. Jiang, J.; Pavlova, N.N.; Zhang, J. Asparagine, a critical limiting metabolite during glutamine starvation. *Mol. Cell Oncol.* **2018**, *5*, e1441633. [CrossRef]

59. Gadian, D.G. *NMR and Its Applications to Living Systems*; Oxford University Press: New York, NY, USA, 1996.