Journal of Pathology
*J Pathol* February 2022; **256:** 202–213
Published online 4 December 2021 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/path.5828

ORIGINAL PAPER

# Metabolic tumor constitution is superior to tumor regression grading for evaluating response to neoadjuvant therapy of esophageal adenocarcinoma patients

Achim Buck[1†], Verena M Prade[1†], Thomas Kunzke[1], Annette Feuchtinger[1], Dino Kröll[2,3], Marcus Feith[4], Bastian Dislich[5], Benjamin Balluff[6], Rupert Langer[5,7‡*] and Axel Walch[1‡*]

[1] *Research Unit Analytical Pathology, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany*
[2] *Department of Visceral Surgery and Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland*
[3] *Department of Surgery, Campus Charité Mitte/Campus Virchow-Klinikum, Charité-Universitätsmedizin Berlin, Berlin, Germany*
[4] *Department of Surgery, Klinikum rechts der Isar, TUM School of Medicine, Munich, Germany*
[5] *Institute of Pathology, University of Bern, Bern, Switzerland*
[6] *Maastricht Multimodal Molecular Imaging Institute (M4i), Maastricht University, Maastricht, The Netherlands*
[7] *Institute of Clinical Pathology and Molecular Pathology, Kepler University Hospital and Johannes Kepler University, Linz, Austria*

**Correspondence to: A Walch, Research Unit Analytical Pathology, Helmholtz Zentrum München – German Research Center for Environmental Health, Ingolstädter Landstraße 1, Neuherberg, 85764, Germany. E-mail: axel.walch@helmholtz-muenchen.de; or R Langer, Institute of Clinical Pathology and Molecular Pathology, Kepler University Hospital, Krankenhausstraße 9, 4021 Linz, Austria. E-mail: rupert.langer@kepleruniklinikum.at*

[†]These authors contributed equally to this work.
[‡]These authors share last authorship.

## Abstract

**The response to neoadjuvant therapy can vary widely between individual patients. Histopathological tumor regression grading (TRG) is a strong factor for treatment response and survival prognosis of esophageal adenocarcinoma (EAC) patients following neoadjuvant treatment and surgery. However, TRG systems are usually based on the estimation of residual tumor but do not consider stromal or metabolic changes after treatment. Spatial metabolomics analysis is a powerful tool for molecular tissue phenotyping but has not been used so far in the context of neoadjuvant treatment of esophageal cancer. We used imaging mass spectrometry to assess the potential of spatial metabolomics on tumor and stroma tissue for evaluating therapy response of neoadjuvant-treated EAC patients. With an accuracy of 89.7%, the binary classifier trained on spatial tumor metabolite data proved to be superior for stratifying patients when compared with histopathological response assessment, which had an accuracy of 70.5%. Sensitivities and specificities for the poor and favorable survival patient groups ranged from 84.9% to 93.3% using the metabolic classifier and from 62.2% to 78.1% using TRG. The tumor classifier was the only significant prognostic factor (HR 3.38, 95% CI 1.40–8.12, $p = 0.007$) when adjusted for clinicopathological parameters such as TRG (HR 1.01, 95% CI 0.67–1.53, $p = 0.968$) or stromal classifier (HR 1.86, 95% CI 0.81–4.25, $p = 0.143$). The classifier even allowed us to further stratify patients within the TRG1–3 categories. The underlying mechanisms of response to treatment have been figured out through network analysis. In summary, metabolic response evaluation outperformed histopathological response evaluation in our study with regard to prognostic stratification. This finding indicates that the metabolic constitution of the tumor may have a greater impact on patient survival than the quantity of residual tumor cells or the stroma.**



**Keywords:** esophageal cancer; spatial metabolomics; imaging mass spectrometry; patient stratification; metabolic response evaluation; tumor regression grading; artificial intelligence; machine learning

Received 20 July 2021; Revised 4 October 2021; Accepted 28 October 2021

*No conflicts of interest were declared.*

## Introduction

Neoadjuvant chemoradiotherapy or chemotherapy is associated with a significant survival benefit for patients compared with surgery alone and has become the standard of care for most patients with resectable esophageal and gastroesophageal-junction adenocarcinoma [1–4]. Preoperative treatment has the effect of tumor- and nodal-downstaging, which can increase the prospect of complete resection [5,6]. Despite the advantages achieved through multimodal therapy, the outcome of a significant proportion of patients with advanced

esophageal adenocarcinoma (EAC) remains unsatisfactory. Insights into the tumor's constitution and its molecular composition after neoadjuvant treatment are valuable for defining risk groups and in order to improve future therapies for patients with a poor response.

A commonly used method to assess the response to neoadjuvant therapy is histological tumor regression grading (TRG), which, in addition to the presence or absence of lymph node metastases, is an important clinically prognostic indicator for treatment response and survival prognosis [7–11]. With regard to TRG, several systems have been proposed in gastrointestinal malignancies aiming to categorize the extent of regressive changes after cytotoxic treatment by estimating the amount of residual tumor and the degree of therapy-induced fibrosis. Thereby, the regressive tissue changes in resected tumor observed after neoadjuvant therapy can range from complete regression to varying amounts of a vital residual tumor. In addition, it has been suggested that the assessment of the tumor metabolic activity using 18-fluorodeoxyglucose-PET ([$^{18}$F]FDG-PET) may help to further improve risk stratification of patients [12]. However, the use of single determined parameters such as tumor glucose uptake or tumor size can be limited in their accuracy for patient response evaluation [13]. To date, no study has addressed the potential of classifiers based on metabolites and their spatial distribution within the tumor to distinguish patients who would benefit from neoadjuvant EAC treatment from those who would not. Spatial metabolomics has already shown promise as a tool to gain new insights into neoplastic progression to EAC [14], for the assessment of surgical resection margins [15], or for investigating tumor heterogeneity [16].

In the present study, we investigated whether spatial metabolomics analysis could improve the assessment of treatment response compared with the histopathological TRG. High mass resolution matrix-assisted laser desorption/ionization Fourier-transform ion cyclotron resonance imaging mass spectrometry (MALDI FTICR IMS) was used to spatially detect metabolites within tissue sections with high sensitivity and specificity. The metabolite features in tumor and stroma tissue compartments were separately investigated with respect to their relevance for therapy response evaluation of EAC patients and used to train random forest classifiers. In univariate and multivariate survival analyses, we evaluated the metabolic-based stratifications with regard to their prognostic power in comparison to established clinicopathological parameters. In addition, the classifier was confirmed to be specific for neoadjuvant treatment using an independent cohort of primary resected EACs as a reference.

## Materials and methods

### Ethics and patient sample size

This study was conducted on two retrospective patient cohorts of neoadjuvant-treated and primary resected EAC cases collected between 1990 and 2011 at the Institutes of Pathology of the University of Bern (Switzerland) and the Technische Universität München (Germany). Tissue samples of these cohorts had been processed into formalin-fixed and paraffin-embedded (FFPE) tissue microarrays (TMAs) as described previously [17]. In brief, tissue cores were annotated by a pathologist (RL) on FFPE tissue blocks/slides, punched out, and transferred into a TMA format as follows: three cores of the tumor bed from each neoadjuvant-treated and six cores from each primary surgical resected case. Within the neoadjuvant-treated EACs ($n = 144$), 59 patients had been treated with platinum/5-fluorouracil (5-FU) and 85 patients with 45 Gy + platinum/5-FU, followed by esophagectomy after completion of the neoadjuvant therapy. The second cohort comprised patients ($n = 64$) who had undergone primary surgical resection without neoadjuvant chemotherapy or radiation. To take into account that the neoadjuvant-treated EACs were UICC-classified as pT3 and pT4, we selected only these categories from the primary resected EAC cohort for reasons of comparability. Clinicopathological characteristics of the patients, such as (y)pTNM categories and resection status, were determined according to the eighth edition of the UICC/AJCC TNM classification [18]. The resected specimens of neoadjuvant-treated EACs were classified according to Becker TRG [5,7]. The basis was the complete histologic examination of the entire tumor bed. To calculate the specificity, sensitivity, and accuracy of TRG, responders were classified as TRG1a, 1b and non-responders as TRG2, 3. The classification into these groups has previously been demonstrated to be prognostically indicative [19]. Overall survival (OS) was calculated from the day of surgery. The usage of FFPE-embedded tumor material for research was approved by the local ethics commissions (Kantonale Ethikkommission Bern, Switzerland, 200/14 and Medizinische Fakultät of the Technische Universität München, Germany, 2056/08).

### Tissue processing and Fourier-transform ion cyclotron resonance (FTICR) MALDI IMS analysis

Tissue preparation steps for the MALDI IMS analysis were performed as previously described [20,21]. In brief, FFPE TMAs were sectioned at a thickness of 4 µm (Microm, HM340E; Thermo Fisher Scientific, Waltham, MA, USA) and mounted onto indium tin oxide (ITO)-coated glass slides (Bruker Daltonik, Bremen, Germany) pretreated with 1:1 poly-L-lysine (Sigma-Aldrich, Munich, Germany) and 0.1% Nonidet P-40 (Sigma-Aldrich). Prior to MALDI matrix application, FFPE sections were adhered by incubating the slide for 1 h at 70 °C, followed by deparaffinization in xylene (Carl Roth, Karlsruhe, Germany) (2 × 8 min), and air drying. Spray coating of tissue sections with 9-aminoacridine (9-AA) hydrochloride monohydrate (Sigma-Aldrich) matrix at 10 mg/ml in 70% methanol (Sigma-Aldrich) was conducted using a SunCollect™ sprayer (Sunchrom, Friedrichsdorf, Germany). The matrix was applied in eight passes (ascending flow rates

of 10, 20, 30 µl/min for layers 1–3, and of 40 µl/min for layers 4–8), using a line distance of 2 mm and a spray velocity of 900 mm/min.

Metabolites were detected in negative-ion mode on a 7 T Solarix XR FT-ICR mass spectrometer (Bruker Daltonik) equipped with a dual ESI-MALDI source and a Smartbeam-II Nd:YAG (355 nm) laser. Data acquisition parameters were specified in ftmsControl software 2.2 and flexImaging (v. 5.0) (Bruker Daltonik). Mass spectra were acquired covering *m/z* 75–1100. The instrument was calibrated externally with L-arginine (Sigma-Aldrich) in the ESI mode and internally using the 9-AA matrix ion signal (*m/z* 193.0771) as lock mass. The laser operated at a frequency of 1000 Hz using 100 laser shots per pixel with a pixel resolution of 60 µm.

### Immunophenotype-guided imaging mass spectrometry

The SPACiAL workflow was used as previously described to automatically annotate tumor and stroma regions in EAC tissues [16]. In brief, after MALDI IMS analysis, the 9-AA matrix was removed from the TMA tissue sections using 70% ethanol (Carl Roth) for 5 min, followed by immunohistochemical staining of the very same tissue sections. TMAs were stained by immunofluorescence using antibodies specific for Pan-Cytokeratin cocktail Plus [AE1/AE3+8/18], 1:75 (Cat# CM162c; Biocare Medical, Pacheco, CA, USA) and vimentin, 1:500 (Cat# ab92547; Abcam, Cambridge, UK). Signal detection was conducted using fluorescence-labeled secondary antibodies purchased from Thermo Fisher Scientific, including anti-rabbit IgG DyLight 633, 1:100 (Cat# 35563) and anti-mouse IgG Alexa Fluor 750 antibody, 1:200 (Cat# A-21037). Nuclei were identified using Hoechst 33342 (Thermo Fisher Scientific). Regions positive for pan-cytokeratin were defined as tumor. Regions negative for pan-cytokeratin but positive for vimentin were defined as stroma. In total, tumor cells in 144 patient samples and stromal compartments in 140 samples could be annotated in the neoadjuvant EAC cohort using these selection criteria. Fluorescence-stained TMAs were imaged automatically using an AxioScan.Z1 digital slide scanner (Carl Zeiss, Göttingen, Germany) equipped with a 20× magnification objective, and visualized using ZEN 2.3 blue edition software (Carl Zeiss). Images were exported as TIF files.

### Peak picking and metabolite annotation

MALDI IMS data were loaded into SCiLS Lab (v. 2020b Pro, Bruker Daltonik). The mass spectra were root-mean-square normalized, and 6945 picked peaks were exported as imzML files for further data processing and subsequent bioinformatics data analysis using an in-house Python 3 pipeline as described previously [16]. Metabolites were annotated based on accurate mass matching (≤ 10 ppm) of the spectrometric *m/z* values to HMDB and KEGG databases allowing $M - H$, $M - H_2O$, $M + K - 2H$, $M + Na - 2H$, and $M + Cl$ as negative adducts.

### General statistical analysis

Statistical analyses were performed in R [22] and Python 3.7 [23]. Comparisons between groups and categories were performed using the rank-based Mann–Whitney *U*-test and Spearman rank-order correlation for continuous data. The resulting *P* values were adjusted with the Benjamini–Hochberg correction (Python 3.7, StatsModels 0.9.0). Univariate analysis for overall survival encompassed Kaplan–Meier analyses and log-rank tests (R, 'survival' package; Python 3.7, lifelines 0.24.8). The intensity threshold dividing high and low abundances was determined for each molecular compound by calculating the log-rank test of each possible split and ensuring robust results by verifying similar results for nearby intensity splits. Unless stated otherwise, *P* values below 0.05 were considered significant for all statistical tests.

### Random forest classifier

Prior to the training of the random forest classifier, features were filtered based on the significance level that they reached in the log-rank test ($p < 0.01$). The feature selection step is a common procedure to identify relevant candidate metabolites from mass spectrometric data for the learning task, minimizing performance losses of the learning algorithm used [24].

A leave-one-out cross-validation approach was used to predict short- or long-term survival patients (Python 3.7, scikit-learn 0.21.3). The training data sets comprise patients from the short-survival group, defined as uncensored patients who survived for less than 48 months after surgery ($n = 31$), and the long-term survivors, comprising all patients who survived for longer than 48 months ($n = 47$). For the leave-one-out cross-validation approach, 200 random forest classifiers were trained per patient and the final prediction was chosen based on a majority vote for each of the 144 patients. The performance of the classifier was evaluated by calculating the accuracy, as well as the sensitivity and specificity. The sensitivity (true positive rate) measures the classifier's ability to correctly classify positive samples. It is calculated as the proportion of positive samples that are correctly identified. The specificity (true negative rate) measures the classifier's ability to correctly classify negative samples. It is calculated as the proportion of negative samples that are correctly identified. The importance of a feature is computed as the (normalized) total reduction of the criterion brought by that feature (Gini importance).

### Cox proportional hazards regression analysis

Multivariate survival analyses were performed using Cox proportional hazards regression (Python 3.7, lifelines 0.24.8). For the classifier comparison to primary resected surgical EAC tissues, the Cox proportional

hazard model was calculated for individual metabolites using the parameters defined for the neoadjuvant classifier. Hence, the previously calculated optimal intensity thresholds separating short- and long-term survivors were applied on primary resected tissues.

### Spatial correlation metabolic networks

The spatially resolved abundances of all metabolites were used to calculate their spatial correlation using the pairwise Spearman rank-order correlation (Python 3.7, SciPy 1.2.0). The resulting *P* values were adjusted with the Benjamini–Hochberg correction (Python 3.7, StatsModels 0.9.0) and filtered ($p < 0.05$). The resulting correlation coefficients were visualized in metabolic networks using Cytoscape (v. 3.8.0). All networks were visualized using the edge-weighted spring embedded layout and the absolute value of the correlation coefficient.

## Results

### Histopathological tumor regression and clinicopathological characteristics

The histopathological classification of the neoadjuvant-treated patient cohort by TRG according to Becker was as follows: 29 patient-resected specimens showed a complete tumor regression (TRG1a: 20.2%), 39 a subtotal tumor regression (TRG1b: 27.1%), 28 a partial response (TRG2: 19.4%), and 48 patients were without evidence of a treatment effect (TRG3: 33.3%). The clinicopathological characteristics are detailed in supplementary material, Table S1. In univariate analysis, TRG ($p = 0.009$) was found to be significantly indicative for the overall survival (OS) of neoadjuvant-treated patients. Other clinicopathological variables including ypT ($p < 0.0001$), ypN ($p = 0.01$), the presence of distant metastasis ($p = 0.0005$), and incomplete resection status ($p < 0.0001$) were found to have a negative impact on survival. The Kaplan–Meier survival curves of TRG and lymph node metastasis are shown in supplementary material, Figure S1A,B.

### Spatial metabolomics improves the assessment of treatment response compared with tumor regression grading

We used our recently published SPACiAL metabolomics approach of immunophenotype-guided imaging mass spectrometry to identify molecular features with potential prognostic value (Figure 1) [16]. Patient tissue sections were automatically segmented into tumor and stromal tissue compartments by multiplex immunohistochemistry, resulting in 10 775 spectra for the tumor class and 11 240 spectra for the stroma class. Approximately 7000 distinct molecular features were detected in tissues of all patient samples. In total, 495 metabolites were found to be predictors of OS according to annotated tumor cells, and 868 hits were found for the stroma region. There was an overlap of 109 metabolites found in both tissue components, demonstrating that automatic separation of regions of interest led to a distinguishable

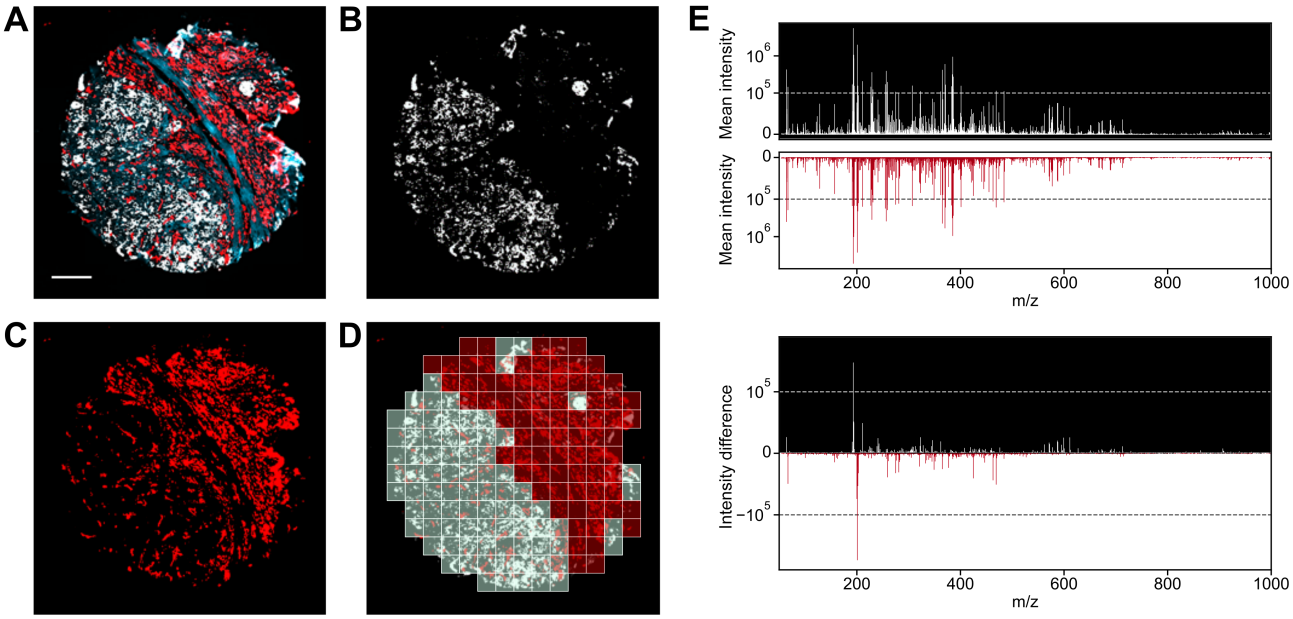
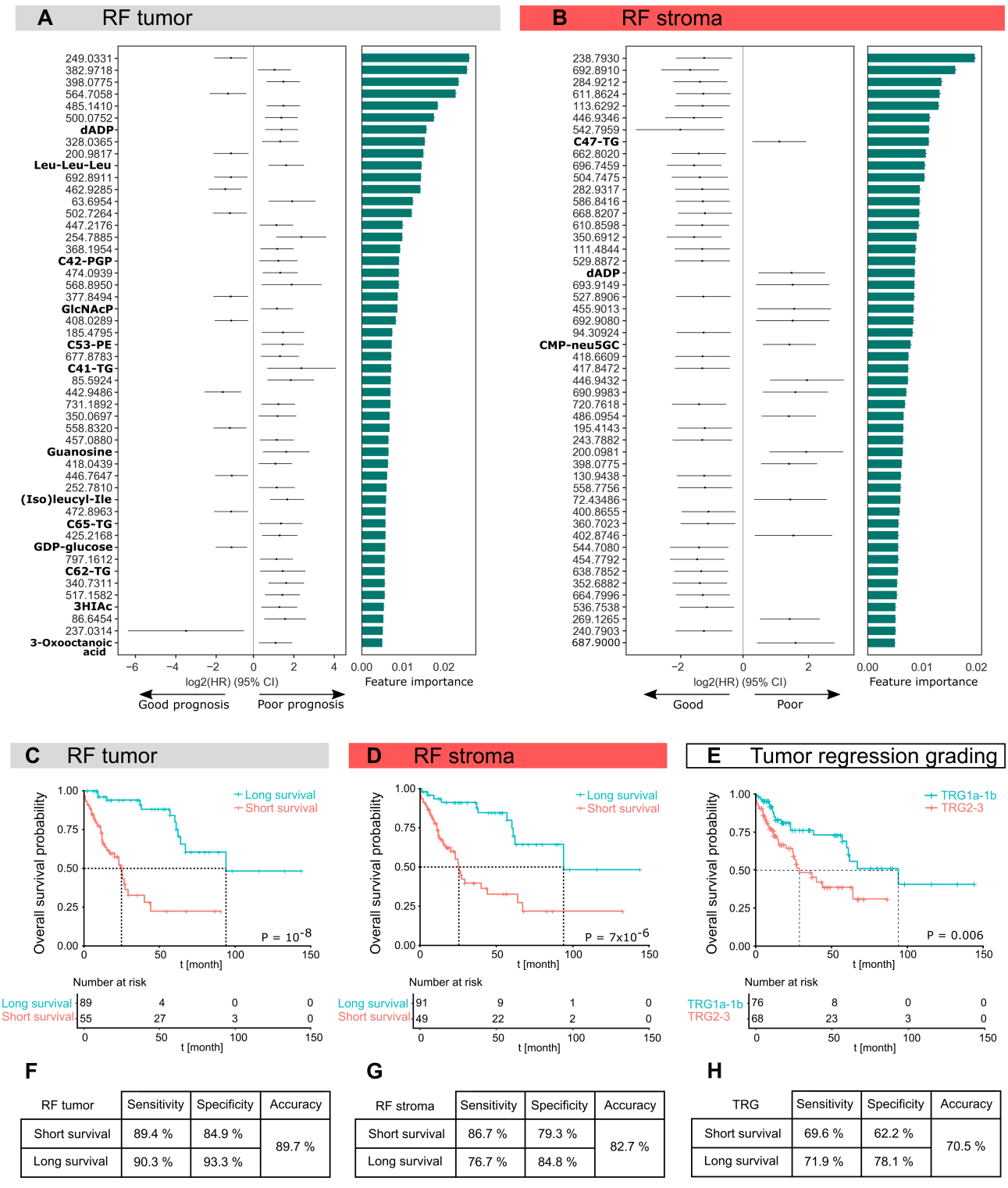



**Figure 1.** Multiplex immunohistochemistry-guided imaging mass spectrometry for metabolic analysis of tumor and tumor stroma. (A) Representative of all three channels of a fluorescence-stained tissue core. White represents the tumor marker (pan-cytokeratin), red the stromal marker (vimentin), and blue the Hoechst channel highlighting the cell nuclei. (B, C) Single channel images of pan-cytokeratin and vimentin used to annotate and separate tumor and stroma cells by fluorescence imaging. (D) Schematic representation of the spatial correlation image analysis (SPACiAL) for extracting mass spectra from regions of interest. (E) Juxtaposition of the overall mean spectrum from tumor (white) and stroma (red) regions, represented as mean intensity and intensity difference. Scale bar: 100 µm.




**F**

| RF tumor | Sensitivity | Specificity | Accuracy |
|---|---|---|---|
| Short survival | 89.4 % | 84.9 % | 89.7 % |
| Long survival | 90.3 % | 93.3 % | |

**G**

| RF stroma | Sensitivity | Specificity | Accuracy |
|---|---|---|---|
| Short survival | 86.7 % | 79.3 % | 82.7 % |
| Long survival | 76.7 % | 84.8 % | |

**H**

| TRG | Sensitivity | Specificity | Accuracy |
|---|---|---|---|
| Short survival | 69.6 % | 62.2 % | 70.5 % |
| Long survival | 71.9 % | 78.1 % | |

**Figure 2.** Random forest (RF) classifiers based on metabolites from tumor and stroma tissue components result in individual risk stratification of patients after neoadjuvant treatment. Forest plots of $\log_2$ hazard ratios with 95% confidence intervals obtained for the best performing classification to separate patients into metabolic responding groups for (A) tumor and (B) tumor stroma. The plots are ordered according to the importance of individual metabolite features. (C, D) Kaplan–Meier survival analyses, as a function of the metabolic subtype of the tumor and stroma classifiers, were used to assess differences in patient overall survival. (E) In comparison, the histopathological tumor regression grading and survival of patients stratified according to responders (TRG1a-1b) and non-responders (TRG2-3). (F, G) Diagnostic performance summarizing the sensitivity, specificity, and accuracy of the tumor and stroma classifiers. (H) Diagnostic performance of TRG. *P* values less than 0.05 were considered significant. 3HIAc, 3-hydroxyisovalerylcarnitine; CMP-neu5GC, CMP-*N*-glycoloylneuraminate; C42-PGP, C42-phosphatidylglycerolphosphate; dADP, deoxyadenosine diphosphate; GDP-glucose, guanosine diphosphate glucose; GlcNAcP, *N*-acetyl-glucosamine phosphate; (Iso)leucyl-Ile, (iso)leucyl-(iso)leucine; Leu-Leu-Leu, trileucine; TG, triglyceride.

Table 1. Multivariate analyses incorporating clinico-morphological parameters into the model demonstrated that the random forest (RF) tumor classifier is a significantly independent factor to predict survival.

| Covariate | HR | 95% CI | *P* value |
|---|---|---|---|
| RF tumor | 3.38 | 1.40–8.12 | **0.007** |
| RF stroma | 1.86 | 0.81–4.25 | 0.143 |
| TRG | 1.01 | 0.67–1.53 | 0.968 |
| ypT | 1.19 | 0.83–1.70 | 0.351 |
| ypN | 1.11 | 0.81–1.52 | 0.521 |
| Distant metastases | 1.47 | 0.69–3.13 | 0.318 |

The value in bold is statistically significant.

number of hits in the molecular output between tumor and stroma. The random forest classifier was then trained separately on metabolite features within tumor and stroma regions. The classifiers with the top 50 metabolites ranked by feature importance and $\log_2$ hazard ratio are presented for tumor in Figure 2A and for stroma in Figure 2B. The top feature (*m/z* 249.0331) in the tumor classifier is 33% more important than the top feature (*m/z* 238.7930) in the stroma classifier. Among the top 50 metabolites in the tumor classifier, 18 metabolites were significantly different between chemotherapy and chemoradiotherapy. In univariate analysis, the metabolic classifiers exhibit significant prognostic power for tumor ($p = 10^{-8}$) and for stroma ($p = 7 \times 10^{-6}$) (Figure 2C, D). In addition, the significance values of individual metabolites were compared with the performance of the classifiers (supplementary material, Figure S2). The survival classifier achieved an accuracy of 89.7% for tumor metabolites and 82.7% for stroma metabolites. The tumor classifier achieved a sensitivity of 89.4% for short survival and 90.3% for long survival and a specificity of 84.9% and 93.3%, respectively (Figure 2F). For the stroma model, the sensitivity was 86.7% with a specificity of 79.3% for patients with short-term survival and the sensitivity was 76.7% with a specificity of 84.8% for patients with long-term survival (Figure 2G). Testing a classification approach that includes metabolites from both tissue compartments, tumor and stroma, did not further improve patient stratification compared with the tumor-only classifier (supplementary material, Figure S3). To evaluate the metabolic classifiers, we used histopathological TRG as a benchmark. In univariate analysis, the Becker TRGs was related to OS (supplementary material, Figure S1A); this becomes even more obvious when TRGs are classified into responders (TRG1a–1b) and non-responders (TRG2–3) ($p = 0.006$) (Figure 2E). According to this classification, TRG achieves an accuracy of 70.5% in survival prognosis (Figure 2H). The sensitivity of TRG classification ranges from 69.6% to 71.9% and the specificity from 62.2% to 78.1% for patients with short-term and long-term survival, respectively. Compared with TRG, the tumor classifier thus achieves a 15–20% higher sensitivity, specificity, and accuracy (Figure 2F,H). In multivariate analysis, the classifier for tumor was the only significant factor for OS (HR 3.38, 95% CI 1.40–8.12, $p = 0.007$) in comparison to TRG (HR 1.01, 95% CI 0.67–1.53, $p = 0.968$), ypT (HR 1.19, 95% CI 0.83–1.70, $p = 0.351$), lymph node metastasis (HR 1.11,

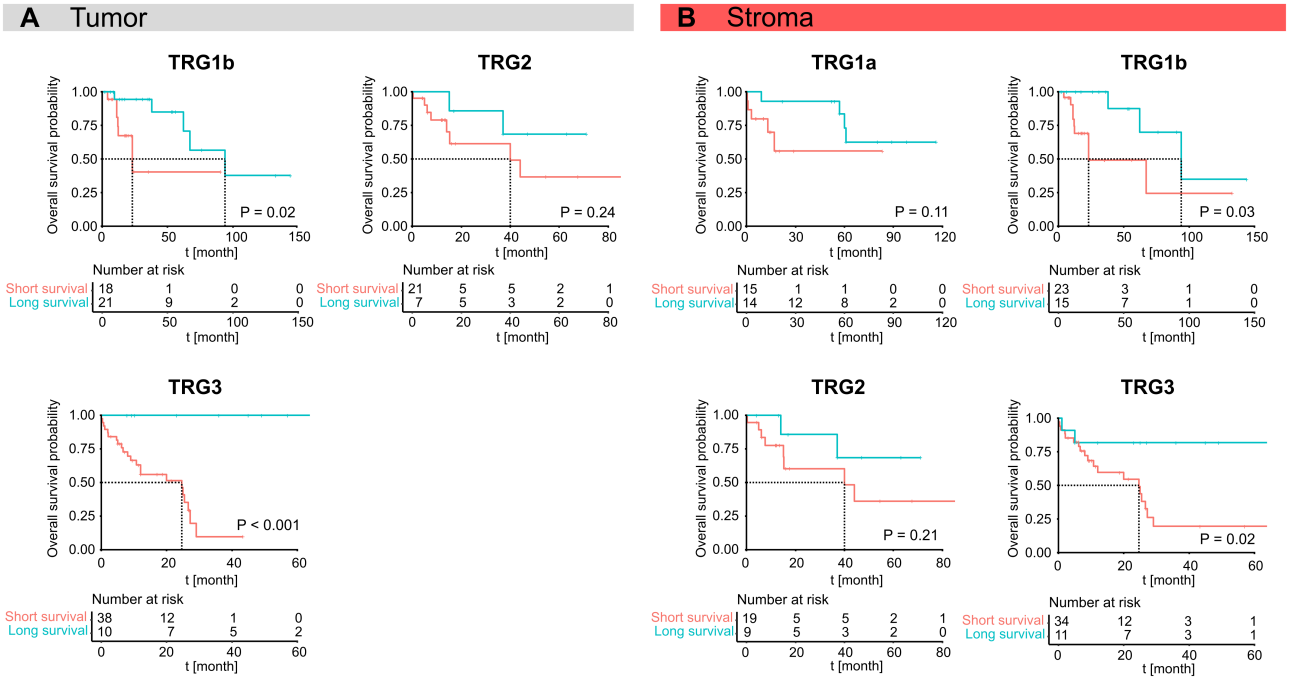



**Figure 3.** Random forest classifiers applied to subgroups of the TRGs enable further patient stratification. (A) Kaplan–Meier analyses significantly distinguished long-term from short-term survivors within TRG1b and TRG3, while TRG2 could not be stratified significantly. Patients with a pathological complete response (TRG1a) were excluded from sub-stratification due to the absence of tumor cells. (B) Stroma classification resulted in a significant stratification of patient survival within TRG1b and TRG3; classification of TRG1a and TRG2 was not significant. *P* values less than 0.05 were considered significant.

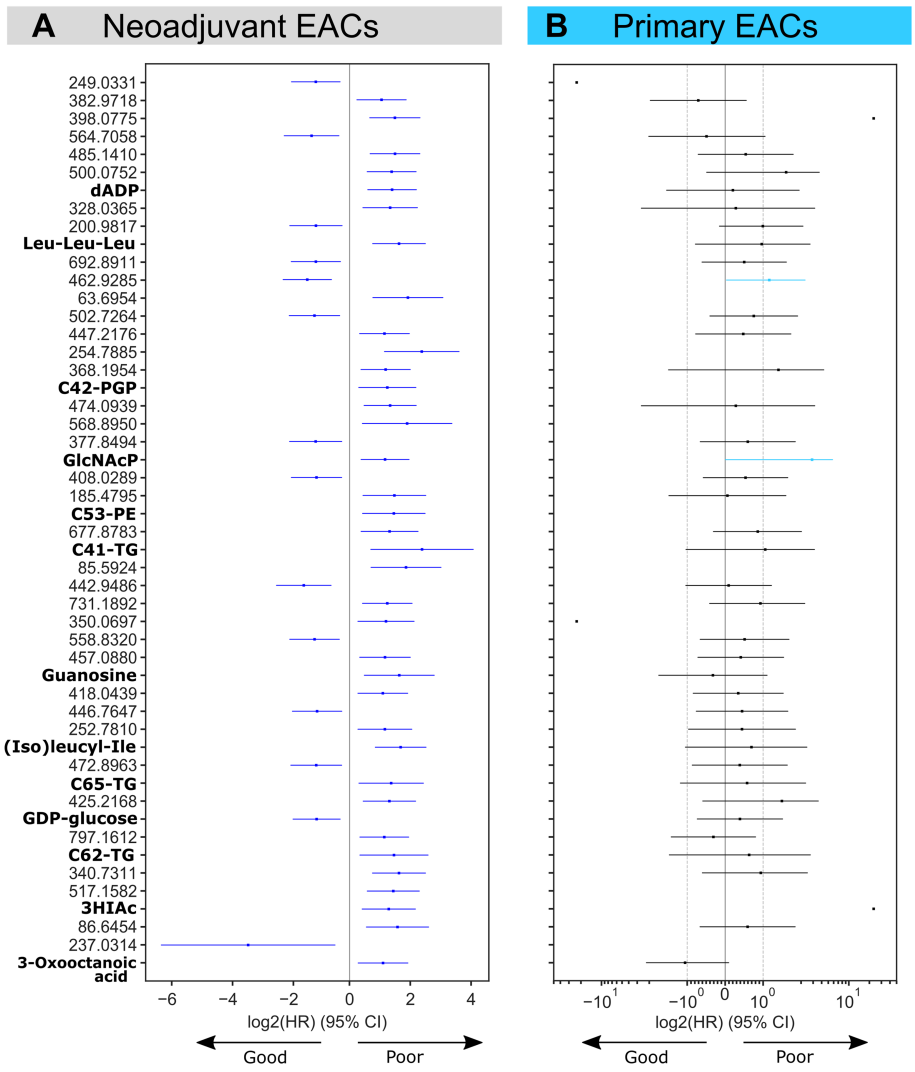



**Figure 4.** Forest plots of $\log_2$ hazard ratios with 95% confidence intervals comparing metabolite levels between neoadjuvant-treated and primary resected EACs. The predictive power of the metabolic features was compared between (A) neoadjuvant and (B) primary resected EACs of patients who had not undergone neoadjuvant treatment. The Cox proportional hazard model was calculated for the top 50 ranked metabolite features on primary EACs. Only two metabolites were significant in primary EACs (highlighted in blue), indicating that most of the metabolic features of the neoadjuvant EAC classifier are related to treatment.

95% CI 0.81–1.52, $p = 0.521$), and distant metastases (HR 1.47, 95% CI 0.69–3.13, $p = 0.318$) as covariables, and was even the stronger factor compared with the stromal classifier (HR 1.86; 95% CI 0.81–4.25, $p = 0.143$) (Table 1).

## Metabolic classifiers of tumor and stroma are even able to stratify categories of TRG

Next, we tested the impact of the tumor and stroma classifiers on patient outcome in relation to the histological response of TRG. Using Kaplan–Meier survival estimates, the tumor random forest classifier was able to further stratify categories of TRG1b ($p = 0.02$) and TRG3 ($p < 0.001$), while TRG2 was not statistically significant ($p = 0.24$) (Figure 3A). Patients with a pathological complete response (TRG1a) were excluded from substratification with the tumor classifier; though stromal changes may occur, tumor cells are no longer present in these cases. Applying the stroma classifier did provide significant differences for OS in TRG1b ($p = 0.03$) and TRG3 ($p = 0.02$), while TRG1a ($p = 0.11$) and TRG2 ($p = 0.21$) were not statistically significant (Figure 3B). This highlights that the metabolic constitution of tumor and stroma has a stronger impact on the treatment response and survival prognosis than the quantity of tumor cells determined by TRG. In addition, the presence of lymph node metastasis is an important prognostic factor related to patient outcome. Therefore, we tested the classifiers' independence with regard to the level of nodal metastasis. The tumor classifier could be used to significantly stratify lymph node status ypN0 ($p < 0.0001$) and ypN1 ($p < 0.001$), while ypN2 and ypN3 were not statistically significant (supplementary material, Figure S4A). By using the stroma classifier, a significant difference could be shown for ypN0 ($p = 0.02$) and ypN1 ($p < 0.001$) but not for ypN2 ($p = 0.61$) (supplementary material, Figure S4B).

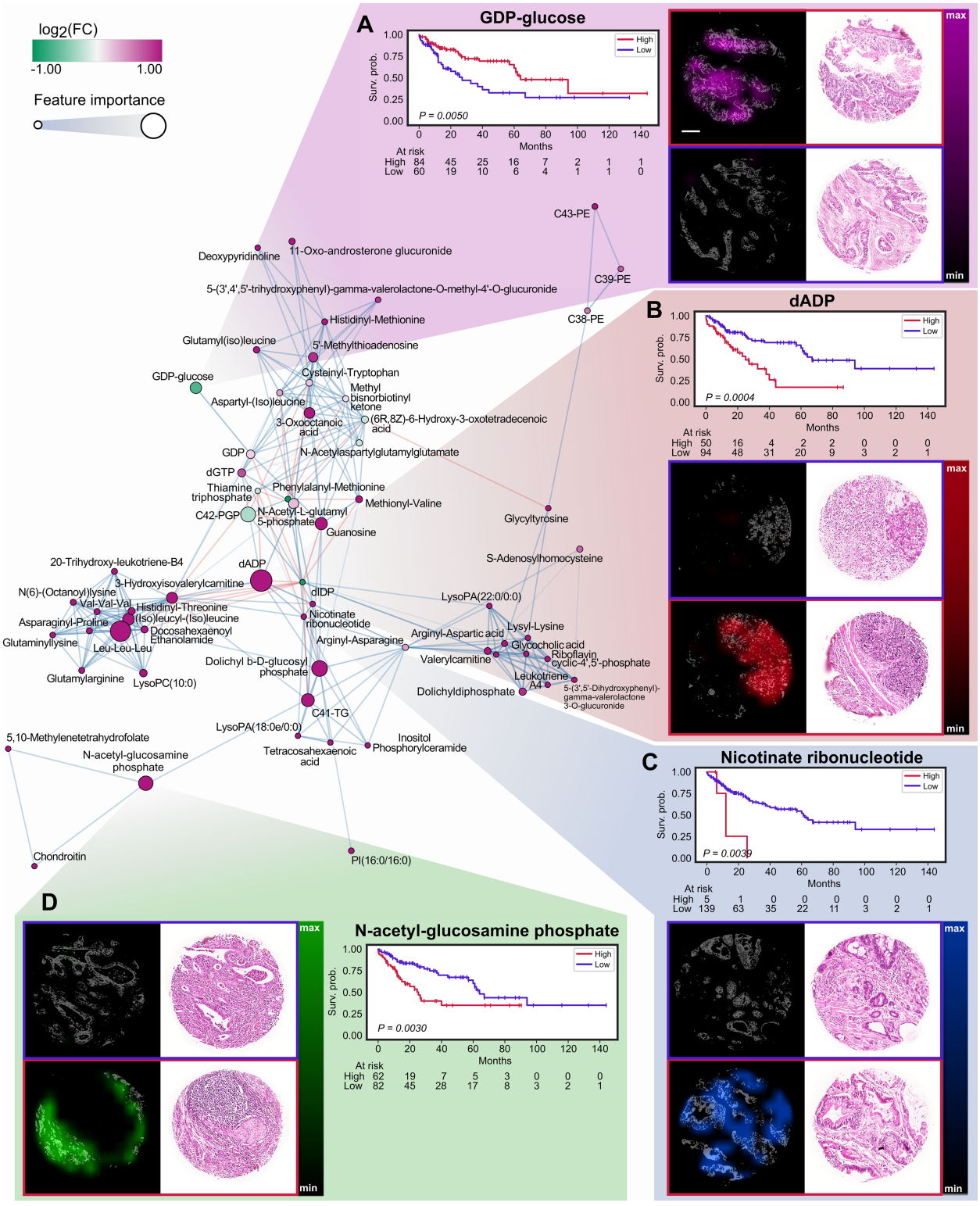



**Figure 5.** Spatial correlation network of metabolites distinguishing between good and poor survival groups in the tumor response evaluation model. Pixel-wise spatial correlations within and between metabolites were calculated and filtered ($p < 0.05$). Node color in the network indicates whether individual metabolites are more abundant in the long-term (green) or short-term survivor (magenta) group. The nodes correlate in size with the feature importance of the classified metabolites. Edges represent positive (blue) and negative (red) spatial correlations between metabolites. Representative images of metabolites from the network are shown with Kaplan–Meier survival analyses for (A) GDP-glucose, (B) dADP, (C) nicotinate ribonucleotide, and (D) *N*-acetyl-glucosamine phosphate. Ion distribution maps reveal specific localization of the metabolite markers in tumor cell regions (white); the corresponding H&E stains of the very same patient tissue core are shown to the right. Patient tissue cores were selected representatively according to their prognostic risk for low mass intensity (image with violet border) or high mass intensity (image with red border). Scale bar: 100 µm. dADP, deoxyadenosine diphosphate; dGTP, deoxyguanosine triphosphate; dIDP, deoxyinosine diphosphate; GDP, guanosine diphosphate; GDP-glucose, guanosine diphosphate glucose; Leu-Leu-Leu, trileucine; LysoPA, lysophosphatidic acid; LysoPC, lysophosphatidylcholine; PE, phosphatidylethanolamine; PGP, phosphatidylglycerophosphate; PI, phosphatidylinositol; TG, triglyceride; Val-Val-Val, trivaline.

### Tumor metabolites are specific for neoadjuvant therapy

In order to validate whether the classifier is specific in terms of treatment, metabolite levels of the neoadjuvant EACs were compared with the levels in primary resected EACs of equivalent stage (Figure 4). We focused on the tumor classifier, due to its stronger effects in patient stratification. In Figure 4B, the calculated hazard ratios of the top 50 stratifying metabolites defined by the neoadjuvant classifier are shown for the primary resected treatment-naïve EAC cohort. There were substantial differences between both cohorts. Only two of the 50 metabolite features, *N*-acetyl-glucosamine phosphate and *m/z* 462.9285, were also significant in primary EAC. The effect of *m/z* 462.9285 changed from good in the neoadjuvant EACs to poor in the primary EACs. Presumably these markers are nonspecific in terms of neoadjuvant treatment, rather than general prognostic markers in EAC biology. This indicates that changes in metabolite levels occur during neoadjuvant therapy.

### Spatial correlations reveal potential functional relationships of metabolites associated with therapy response

To investigate the metabolic characteristics related to treatment response, we performed a metabolic network analysis evaluating the co-localization pattern of the metabolites in the neoadjuvant tumor evaluation model. The spatial correlation network illustrated in Figure 5 reveals the correlation of functionally interconnected metabolites. Five dense clusters within the networks indicate a stronger spatial correlation of the corresponding metabolites and thus a higher dependency between their amounts. Altered metabolism within the clusters particularly includes changes in amino acids and analogues (21.5%), lipid biogenesis (19.8%), nucleosides and nucleotides (6.6%), and carbohydrates and carbohydrate conjugates (4.1%). In the network, the metabolite importance in the classifier is reflected by node size, while the node color indicates whether individual metabolites are more abundant in the long-term or short-term survivor group. Accordingly, GDP-glucose is more important than GDP, and they correlate negatively ($r_s = -0.0177$). High mass intensity of the metabolite in tumor cells is significantly associated with long OS ($p = 0.005$), which is also shown in representative images of GDP-glucose distribution within tissues (Figure 5A). Likewise, the glycerophospholipid C42-PGP showed good prognosis if the metabolite was abundantly detected in the tumor ($p = 0.0082$). Conversely, the nucleotide dADP was associated at low signal intensity with good prognosis and at high intensity with poor prognosis ($p = 0.0004$) (Figure 5B). With a high feature weighting in the classifier, dADP correlates to *N*-acetyl-L-glutamyl 5-phosphate ($r_s = 0.528$), nicotinate ribonucleotide ($r_s = 0.447$), dIDP ($r_s = 0.430$), 3-hydroxyisovalerylcarnitine ($r_s = 0.337$), and thiamine triphosphate ($r_s = -0.312$), while correlations to dolichyl b-D-glucosyl phosphate ($r_s = 0.256$), dGTP ($r_s = 0.213$), histidinyl-threonine ($r_s = 0.210$), and arginyl-asparagine ($r_s = 0.208$) were found to be less pronounced. In addition, many dipeptide and tripeptide profiles, such as trileucine, (iso)leucyl-(iso)leucine, or *N*-acetylaspartylglutamylglutamate, contributed to the separation of the long and short survival groups. Besides, nicotinate ribonucleotide, which along with dADP correlates with ten other metabolites, correlates with good patient prognosis when the mass signal intensity is low ($p = 0.0039$) (Figure 5C). A small cluster was formed by *N*-acetyl-glucosamine phosphate ($p = 0.003$) (Figure 5D) and C41-triglyceride ($r_s = 0.216$), chondroitin ($r_s = 0.219$), and 5,10-methylene-tetrahydrofolate ($r_s = 0.219$).

## Discussion

In the present study, we used tumor metabolism to assess the treatment response of patients with locally advanced adenocarcinoma of the esophagus. We found that in neoadjuvant-treated EACs, the metabolic response is the stronger arbiter of patient outcome when compared with the histopathological TRG.

The metabolic classifiers were primarily compared with the pathological regression according to the Becker TRG, which takes the quantity of tumor cells into account for classification. For the patient stratification into short- and long-term survival groups, metabolic response based on changes in tumor cells exhibited superior results in accuracy, sensitivity, and specificity compared with TRG (Figure 2F,H). This result suggests that not the tumor cell count but rather changes in the tumor's metabolic constitution determine the response of neoadjuvant-treated EAC patients. However, some effects could also be derived from the metabolic state of tumor stroma, although these effects were more pronounced in tumor, as shown by multivariate analysis. The finding that the metabolism of both – the tumor and stromal tissue compartments – provides valuable information about treatment response is consistent with the concepts of regression grading: the TRGs that primarily refer to residual tumor cells as a reference for response evaluation to therapy, such as the Becker system [5,25], and TRGs that additionally incorporate the extracellular matrix (e.g. the Mandard or Dworak systems) [26–28]. It should be noted that the biochemical characteristics of tumor and stroma can be assessed by spatial metabolomics analysis; however, the qualitative assessment of the tumor stroma in particular remains limited by histomorphological TRG. Importantly, using the metabolic classifiers, it was even possible to further stratify patients into different prognostic groups that could not be stratified by TRG. These extended risk stratification in the individual categories of TRG, and nodal status could be achieved globally for tumor and stroma (Figure 3 and supplementary material, Figure S4), demonstrating the effectiveness of the metabolic response evaluation.

A particular feature of our study is that the metabolites of the classifier were tested to be specific for treatment by comparing metabolite levels to a primary resected and treatment-naïve EAC cohort. The majority of the classifier's metabolites in tumor changed significantly in their abundance under therapy. This is consistent with previous studies reporting differences in the expression of molecular and diagnostic markers after neoadjuvant therapy [29,30]. Consequently, the identified metabolites strongly determine patient outcome in terms of treatment, while the classifier's features were mainly nonspecific for patients who underwent primary surgery (Figure 4). These results are remarkable considering that the alternative molecular analysis from small tumor tissue cores of TMAs led to these superior results in patient classification.

To gain insight into the molecules' processes and events that play a role in tumor cells, and which are related to patients' outcome, we performed a metabolic network analysis determining the spatial correlation between metabolites. Altered metabolites between the two survival groups comprise different biomolecular compounds such as amino acids. These are involved in various pathways contributing to cancer cell growth and survival [31], but not much is known about the functional properties of the di- and tri-peptides that had an impact on patient survival in our study. Di- and tripeptides have been suggested as potential disease biomarkers for cancer and have become increasingly the subject of metabolomic studies [32,33]. For instance, the tripeptide *N*-acetylaspartylglutamylglutamate is produced together with the dipeptide *N*-acetyl-aspartylglutamate in cells by the same synthetase [34]. The latter metabolite was unveiled as a glutamate reservoir in diverse cancers [35], while the precise function of *N*-acetylaspartylglutamylglutamate remains to be clarified. Most cancer cells, including esophageal cancer cells but also platinum-treated lung cancer cells, show a strong dependence on an exogenous supply of glutamine, a precursor of glutamate, and cells respond very sensitively to glutamine deprivation [36,37]. Thus, *N*-acetylaspartylglutamylglutamate might be able to store glutamate for use when its production from other sources is limited. Alongside their function as substrates for protein synthesis, amino acids also play a role in energy generation, maintaining cellular redox homoeostasis and driving the synthesis of nucleic acid.

Major changes in nucleoside and nucleotide metabolism have been linked to patient survival. Typically, cancer cells have deactivated crucial DNA damage response signaling routes and often rewire their metabolism and energy production networks [38,39]. Platinum-based chemotherapies and 5-FU are routinely used for neoadjuvant treatment and cause significant changes in DNA metabolism, either by interference or by chemical reaction with DNA [40,41]. Transcriptomics and metabolomics studies using different tumor and non-tumor cell lines uncovered a strong correlation between platinum sensitivity and pathways involved in nucleotide metabolism by salvage or *de novo* synthesis [42,43]. The changes in metabolism, exemplified for the nucleotides dADP, GDP-glucose, and nicotinate ribonucleotide (Figure 5), can be interpreted either as the resistance of chemotherapeutic agents or as an alteration in metabolism occurring in response to DNA repair. Moreover, we found the metabolite 5,10-methylenetetrahydrofolate to be prognostic for EAC. It has been shown that there is a correlation between the activity of methylenetetrahydrofolate reductase (MTHFR), an inhibitor of thymidylate synthase, and the response of tumor cells to the chemotherapeutic agent 5-FU [44]. It has been suggested that intracellular concentrations of 5,10-methylenetetrahydrofolate can increase the cytotoxic effect of 5-FU by the formation of a ternary complex that inhibits the activity of thymidylate synthase and suppresses DNA synthesis. An improved response to 5-FU-based chemotherapy was observed in patients with advanced colorectal cancer bearing a polymorphic MTHFR allele [45], while in contrast, low MTHFR expression levels in EACs showed a poor response to treatment [46].

Our classifier is a starting point that warrants further validation in other independent patient cohorts. Importantly, the patient tissue specimens investigated in this study were routine pathological FFPE resection specimens processed for diagnostic purposes to macroscopically and histologically assess the effects of neoadjuvant therapy. The integrity of the tissue is maintained during metabolic measurement, allowing further development of the assay in combination with histopathology. The technical application of MALDI IMS is already available in an increasing number of pathological institutions. It should be noted that before the method can be applied in laboratories in a clinical setting, consensus is needed in the standardization of procedures and protocols, which is currently the subject of ongoing research [47–49]. In our analyses, we included patients treated with neoadjuvant chemotherapy and chemoradiotherapy. It is conceivable that patient stratification according to treatment regimen could further improve classification. In conclusion, we have demonstrated that our metabolite-based classifier identified from EAC tumor samples of patients treated with neoadjuvant chemoradiotherapy provides valuable information on treatment response and patients' survival independently of TRG and lymph node status. Tumor response evaluation by TRG is commonly used to determine the treatment success of neoadjuvant therapy in EAC patients. We have demonstrated that our metabolic classifier can likewise be used to measure tumor response and may even outperform conventional histopathologic estimation of tumor regression with regard to prognostic stratification of patients.

## Acknowledgements

This study was funded by the European Union (ERA NET: TRANSCAN 2) granted to AW and BB. AW acknowledges the support of the Deutsche Forschungsgemeinschaft (Grant No. SFB 824TP C04). Funding was

provided through the Impulse and Networking Fund of the Helmholtz Association and the Helmholtz Zentrum München (Helmholtz Enterprise-2018-6) to AB. We are grateful to Ulrike Buchholz, Claudia-Mareike Pflüger, Andreas Voss, Cristina Hübner Freitas, and Elenore Samson for their excellent technical assistance.

## Author contributions statement

AB, VMP, RL and AW contributed to conception and design and to manuscript writing. AB and VMP contributed to data acquisition, analysis, visualization, and interpretation. TK, AF and BB contributed to data analyses and bioinformatics assistance. DK, MF, BD and RL contributed to patient characterization and to provision of patient tissue and data. All the authors reviewed and approved the manuscript.

## Data availability statement

The data that support the findings of this study are available from the corresponding authors (AW and RL) on request.

## References

1. Cunningham D, Stenning SP, Smyth EC, *et al*. Peri-operative chemotherapy with or without bevacizumab in operable oesophagogastric adenocarcinoma (UK Medical Research Council ST03): primary analysis results of a multicentre, open-label, randomised phase 2–3 trial. *Lancet Oncol* 2017; **18:** 357–370.
2. van Hagen P, Hulshof MC, van Lanschot JJ, *et al*. Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med* 2012; **366:** 2074–2084.
3. Ychou M, Boige V, Pignon JP, *et al*. Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: an FNCLCC and FFCD multicenter phase III trial. *J Clin Oncol* 2011; **29:** 1715–1721.
4. Al-Batran SE, Homann N, Pauligk C, *et al*. Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. *Lancet* 2019; **393:** 1948–1957.
5. Becker K, Mueller JD, Schulmacher C, *et al*. Histomorphology and grading of regression in gastric carcinoma treated with neoadjuvant chemotherapy. *Cancer* 2003; **98:** 1521–1530.
6. Davies AR, Myoteri D, Zylstra J, *et al*. Lymph node regression and survival following neoadjuvant chemotherapy in oesophageal adenocarcinoma. *Br J Surg* 2018; **105:** 1639–1649.
7. Becker K, Reim D, Novotny A, *et al*. Proposal for a multifactorial prognostic score that accurately classifies 3 groups of gastric carcinoma patients with different outcomes after neoadjuvant chemotherapy and surgery. *Ann Surg* 2012; **256:** 1002–1007.
8. Koh YW, Park YS, Ryu MH, *et al*. Postoperative nodal status and diffuse-type histology are independent prognostic factors in resectable advanced gastric carcinomas after preoperative chemotherapy. *Am J Surg Pathol* 2013; **37:** 1022–1029.
9. Fokas E, Liersch T, Fietkau R, *et al*. Tumor regression grading after preoperative chemoradiotherapy for locally advanced rectal carcinoma revisited: updated results of the CAO/ARO/AIO-94 trial. *J Clin Oncol* 2014; **32:** 1554–1562.
10. Davarzani N, Hutchins GGA, West NP, *et al*. Prognostic value of pathological lymph node status and primary tumour regression grading following neoadjuvant chemotherapy – results from the MRC OE02 oesophageal cancer trial. *Histopathology* 2018; **72:** 1180–1188.
11. Smyth EC, Fassan M, Cunningham D, *et al*. Effect of pathologic tumor response and nodal status on survival in the Medical Research Council Adjuvant Gastric Infusional Chemotherapy trial. *J Clin Oncol* 2016; **34:** 2721–2727.
12. Lordick F, Ott K, Krause BDJ, *et al*. PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial. *Lancet Oncol* 2007; **8:** 797–805.
13. Kwee RM. Prediction of tumor response to neoadjuvant therapy in patients with esophageal cancer with use of $^{18}$F FDG PET: a systematic review. *Radiology* 2010; **254:** 707–717.
14. Abbassi-Ghadi N, Antonowicz SS, McKenzie JS, *et al*. *De novo* lipogenesis alters the phospholipidome of esophageal adenocarcinoma. *Cancer Res* 2020; **80:** 2764–2774.
15. Eberlin LS, Tibshirani RJ, Zhang J, *et al*. Molecular assessment of surgical-resection margins of gastric cancer by mass-spectrometric imaging. *Proc Natl Acad Sci U S A* 2014; **111:** 2436–2441.
16. Prade VM, Kunzke T, Feuchtinger A, *et al*. *De novo* discovery of metabolic heterogeneity with immunophenotype-guided imaging mass spectrometry. *Mol Metab* 2020; **36:** 100953.
17. Galván JA, Wiprächtiger J, Slotta-Huspenina J, *et al*. Immunohistochemical analysis of the expression of cancer-associated fibroblast markers in esophageal cancer with and without neoadjuvant therapy. *Virchows Arch* 2020; **476:** 725–734.
18. Brierley JD, Gospodarowicz MK, Wittekind C. *TNM Classification of Malignant Tumours*. John Wiley & Sons: New York, 2017.
19. Becker K, Langer R, Reim D, *et al*. Significance of histopathological tumor regression after neoadjuvant chemotherapy in gastric adenocarcinomas: a summary of 480 cases. *Ann Surg* 2011; **253:** 934–939.
20. Buck A, Ly A, Balluff B, *et al*. High-resolution MALDI-FT-ICR MS imaging for the analysis of metabolites from formalin-fixed, paraffin-embedded clinical tissue samples. *J Pathol* 2015; **237:** 123–132.
21. Ly A, Buck A, Balluff B, *et al*. High-mass-resolution MALDI mass spectrometry imaging of metabolites from formalin-fixed paraffin-embedded tissue. *Nat Protoc* 2016; **11:** 1428–1443.
22. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, 2021.
23. Rossum GV, Drake FL. *Python 3 Reference Manual*. CreateSpace: Scotts Valley, CA, 2009.
24. Liebal UW, Phan ANT, Sudhakar M, *et al*. Machine learning applications for mass spectrometry-based metabolomics. *Metabolites* 2020; **10:** 243.
25. Japan Esophageal Society. Japanese Classification of Esophageal Cancer, 11th Edition: part I. *Esophagus* 2017; **14:** 1–36.
26. Mandard AM, Dalibard F, Mandard JC, *et al*. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer* 1994; **73:** 2680–2686.
27. Dworak O, Keilholz L, Hoffmann A. Pathological features of rectal cancer after preoperative radiochemotherapy. *Int J Colorectal Dis* 1997; **12:** 19–23.
28. Amin MB, Greene FL, Edge SB, *et al*. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin* 2017; **67:** 93–99.
29. van de Ven S, Smit VT, Dekker TJ, *et al*. Discordances in ER, PR and HER2 receptors after neoadjuvant chemotherapy in breast cancer. *Cancer Treat Rev* 2011; **37:** 422–430.

30. Lee HC, Ko H, Seol H, *et al*. Expression of immunohistochemical markers before and after neoadjuvant chemotherapy in breast carcinoma, and their use as predictors of response. *J Breast Cancer* 2013; **16:** 395–403.
31. Vettore L, Westbrook RL, Tennant DA. New aspects of amino acid metabolism in cancer. *Br J Cancer* 2020; **122:** 150–156.
32. Ozawa H, Hirayama A, Shoji F, *et al*. Comprehensive dipeptide analysis revealed cancer-specific profile in the liver of patients with hepatocellular carcinoma and hepatitis. *Metabolites* 2020; **10:** 442.
33. Soga T, Sugimoto M, Honma M, *et al*. Serum metabolomics reveals γ-glutamyl dipeptides as biomarkers for discrimination among different forms of liver disease. *J Hepatol* 2011; **55:** 896–905.
34. Lodder-Gadaczek J, Becker I, Gieselmann V, *et al*. *N*-Acetylaspartylglutamate synthetase II synthesizes *N*-acetylaspartylglutamylglutamate. *J Biol Chem* 2011; **286:** 16693–16706.
35. Nguyen T, Kirsch BJ, Asaka R, *et al*. Uncovering the role of *N*-acetylaspartyl-glutamate as a glutamate reservoir in cancer. *Cell Rep* 2019; **27:** 491–501.e6.
36. Qie S, Yoshida A, Parnham S, *et al*. Targeting glutamine-addiction and overcoming CDK4/6 inhibitor resistance in human esophageal squamous cell carcinoma. *Nat Commun* 2019; **10:** 1296.
37. Wangpaichitr M, Wu C, Li YY, *et al*. Exploiting ROS and metabolic differences to kill cisplatin resistant lung cancer. *Oncotarget* 2017; **8:** 49275–49292.
38. Dang CV. Links between metabolism and cancer. *Genes Dev* 2012; **26:** 877–890.
39. Lord CJ, Ashworth A. The DNA damage response and cancer therapy. *Nature* 2012; **481:** 287–294.
40. von Stechow L, Ruiz-Aracama A, van de Water B, *et al*. Identification of cisplatin-regulated metabolic pathways in pluripotent stem cells. *PLoS One* 2013; **8:** e76476.
41. Cheung-Ong K, Giaever G, Nislow C. DNA-damaging agents in cancer chemotherapy: serendipity and chemical biology. *Chem Biol* 2013; **20:** 648–659.
42. Cavill R, Kamburov A, Ellis JK, *et al*. Consensus-phenotype integration of transcriptomic and metabolomic data implies a role for metabolism in the chemosensitivity of tumour cells. *PLoS Comput Biol* 2011; **7:** e1001113.
43. Kowalski D, Pendyala L, Daignan-Fornier B, *et al*. Dysregulation of purine nucleotide biosynthesis pathways modulates cisplatin cytotoxicity in *Saccharomyces cerevisiae*. *Mol Pharmacol* 2008; **74:** 1092–1100.
44. Sohn KJ, Croxford R, Yates Z, *et al*. Effect of the methylenetetrahydrofolate reductase C677T polymorphism on chemosensitivity of colon and breast cancer cells to 5-fluorouracil and methotrexate. *J Natl Cancer Inst* 2004; **96:** 134–144.
45. Cohen V, Panet-Raymond V, Sabbaghian N, *et al*. Methylenetetrahydrofolate reductase polymorphism in advanced colorectal cancer: a novel genomic predictor of clinical response to fluoropyrimidine-based chemotherapy. *Clin Cancer Res* 2003; **9:** 1611–1615.
46. Langer R, Specht K, Becker K, *et al*. Association of pretherapeutic expression of chemotherapy-related genes with response to neoadjuvant chemotherapy in Barrett carcinoma. *Clin Cancer Res* 2005; **11:** 7462–7469.
47. Buck A, Heijs B, Beine B, *et al*. Round robin study of formalin-fixed paraffin-embedded tissues in mass spectrometry imaging. *Anal Bioanal Chem* 2018; **410:** 5969–5980.
48. Ly A, Longuespée R, Casadonte R, *et al*. Site-to-site reproducibility and spatial resolution in MALDI-MSI of peptides from formalin-fixed paraffin-embedded samples. *Proteomics Clin Appl* 2019; **13:** e1800029.
49. Schwamborn K, Kriegsmann M, Weichert W. MALDI imaging mass spectrometry – from bench to bedside. *Biochim Biophys Acta Proteins Proteom* 2017; **1865:** 776–783.

## SUPPLEMENTARY MATERIAL ONLINE

**Figure S1.** Overall survival according to tumor regression grade and lymph node status

**Figure S2.** Significance values of individual metabolites compared with the performance of the classifiers

**Figure S3.** Random forest classifier including metabolites from both tissue compartments, tumor and stroma, results in risk stratification of patients after neoadjuvant treatment

**Figure S4.** Tumor and stroma random forest classifiers applied to subgroups of the ypN enable patient stratification

**Table S1.** Clinicopathological variables of 144 patients with EAC who underwent neoadjuvant chemoradiotherapy followed by esophagectomy and 64 patients with primary resected EAC