Spatially resolved metabolic phenotyping of breast cancer by desorption electrospray ionization mass spectrometry

Sabine Guenther¹, Laura J Muirhead², Abigail VM Speller¹, Ottmar Golf¹, Nicole Strittmatter¹, Rathiramakrishnan³, Robert D Goldin³, Emrys A Jones¹, Kirill Veselkov¹, Ara Darzi², Zoltan Takats¹

¹ Section of Computational and Systems Medicine, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, UK
² Section of Biosurgery and Surgical Technology, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, UK
³ Centre for Pathology, Department of Medicine, Imperial College London, UK

Keywords: breast cancer, mass spectrometry imaging, DESI, phenotyping, metabonomics

Running title: Breast cancer diagnosis and characterization by DESI MSI

Precis: In evaluating a new mass spectrometry-based tool to analyze breast tumor tissues, this study shows how it can rapidly infer tumor grade and hormone receptor status from the metabolic profile of fresh frozen sections of resected specimens.

Financial support: European Research Council under the Starting Grant scheme (contract no. 210356), European Commission FP7 Intelligent Surgical Device project (contract no. 3054940) and NIHR Biomedical Research Centre (Imperial College)

Corresponding Author: Prof. Zoltan Takats, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, 371 SAF Building, South Kensington Campus, SW7 2AZ London, United Kingdom. Tel: +44 0207 5942760
email: z.takats@imperial.ac.uk

The authors disclose no potential conflicts of interest.

Word count: 4995 Figures: 4 Tables: 3

Supplemental material: 1073 words, 3 Figures, 1 table.
Abstract

Breast cancer is a heterogeneous disease characterized by varying response to therapeutic agents and significant differences in long-term survival. Thus there remains an unmet need for early diagnostic and prognostic tools and improved histological characterization for more accurate disease stratification and personalized therapeutic intervention. This study evaluated a comprehensive metabolic phenotyping method in breast cancer tissue that uses desorption electrospray ionization mass spectrometry imaging (DESI MSI), both as a novel diagnostic tool and as a method to further characterize metabolic changes in breast cancer tissue and the tumor microenvironment. In this prospective single centre study, 126 intra-operative tissue biopsies from tumor and tumor bed from 50 patients undergoing surgical resections were subject to DESI MSI. Global DESI MSI models were able to distinguish adipose, stromal and glandular tissue based on their metabolomic fingerprint. Tumor tissue and tumor-associated stroma showed evident changes in their fatty acid and phospholipid composition compared to normal glandular and stromal tissue. Diagnosis of breast cancer was achieved with an accuracy of 98.2% based on DESI MSI data (PPV 0.96, NVP 1, specificity 0.96, sensitivity 1). In the tumor group, correlation between metabolomic profile and tumor grade/hormone receptor status was found. Overall classification accuracy was 87.7% (PPV 0.92, NPV 0.9, specificity 0.9, sensitivity 0.92). These results demonstrate that DESI MSI may be a valuable tool in the improved diagnosis of breast cancer in the future. The identified tumor-associated metabolic changes support theories of de novo lipogenesis in tumor tissue and the role of stroma tissue in tumor growth and development and overall disease prognosis.

Introduction

Breast cancer is the most common cancer in women; approximately 12.3% of women in
the United States will be diagnosed with breast cancer at some point in their lifetime. It is a heterogeneous disease with disparate patient outcomes dependant on multiple factors, including: tumor size, nodal status, histological grade and type, peritumoral vascular invasion, HER2 and hormone receptor (HR) status. A proportion of breast cancers will benefit from systemic chemotherapy, endocrine and HER2 based therapies. HR and HER2 receptor expression are strongly predictive biomarkers for the response to targeted therapies. Results of these tests alone are key determinants of which adjuvant treatment the patient receives. There is, however, evidence of variability in assay results with discordance between laboratories, as there is no gold standard IHC assay available (1). Furthermore, these traditional pathological prognostic categories are currently not sufficient to fully elucidate which patients will see the greatest benefit from these treatments. As such, further research into improving the management of breast cancer by early diagnosis, accurate characterization of disease, focused, personalized treatment and risk stratification is essential. Significant progress has been made in the development of tissue based tests which aim to predict risk of recurrence in patients with early stage breast cancer and DCIS. These include gene expression profiling assays such as Oncotype Dx (2), MammaPrint (3), PAM50 (4) and expanded immunohistochemistry tests such as IHC4 (5), and Mammostrat (6).

Both genetic/epigenetic (7-9), and metabolic (10) changes have been demonstrated in the histologically normal tissue surrounding invasive breast cancer compared to distant tissue, or that which has been excised for completely benign disease. These changes support the still incompletely understood theory of field cancerization. There is growing evidence that peritumoral microenvironmental changes may represent underlying mechanisms that affect tumor cell initiation, progression and disease prognosis (11) and may affect surgical margins and risk of recurrence. Dynamic Contrast Enhanced Magnetic Resonance Imaging has been used to identify this tumor associated stroma (12).
Over the last decade, mass spectrometry imaging (MSI) has been increasingly applied to investigate the spatial distribution of biomolecules in tissue sections (13-15). Its potential to reveal the distribution of hundreds of molecules in a single measurement - without prior derivatization - is especially useful for the clinico-pathological analysis of cancer (16-18). Desorption electrospray ionization (DESI) MSI allows the detection of small metabolites and complex lipids within a tissue sample (19, 20). DESI MSI is not per se a quantitative method, but allows for relative quantification when the same tissue type is compared. It may be applied to fresh or frozen tissue sections without the need for further sample preparation, and as such may be readily incorporated into the operating theater and/or histopathological workflow as recently demonstrated for the intraoperative determination of margin status in brain tumor surgery (21-23).

This study applied DESI MSI to the analysis of breast cancer to determine whether this technique can provide further information on the metabolic changes in breast tumor tissue and peritumoral tissue which may provide novel biomarkers for stratification and prognostication of the disease and improve our understanding of tumor/microenvironment metabolomics and field cancerization. Moreover, we investigate the association between the metabolomic profile of the tissue and clinical factors such as cancer type, grade and hormone receptor status.

**Materials and Methods**

**Clinical study and patient recruitment**

This was a single centre, prospective observational study performed at Charing Cross Hospital, London (REC 11/LO/0686). Female patients undergoing breast surgery for benign and malignant disease were recruited on the day of surgery having previously agreed to contribute tissue to this study, which was approved by the local Ethical Review Board. Patients who had undergone neoadjuvant therapy were included, and this was
documented in the corresponding clinical dataset, as was: medical history, drug history, menopausal state, ethnicity, smoking, alcohol status and lifetime use of hormone therapy such as the Oral Contraceptive Pill or Hormone Replacement Therapy. Intra-operative biopsies were taken from the lesion (28 patients, 28 samples) and tumor bed (22 patients, 98 samples), flash frozen and stored at -80°C. The ‘tumor bed’ in this context is used to describe healthy tissue in the cavity beyond the surgical resection margin. Tissue was sampled intra-operatively using a disposable core biopsy gun (Tru-Core II 14G x 10 cm). One biopsy sample was taken from each tumor and 4-5 biopsy samples were taken from the tumor bed; at 3, 6, 9 and 12 o’clock and, from the posterior/deep cavity wall where possible. Demographic and clinicopathological characteristics of the patient cohort are summarized in Table 1.

**Sample preparation and Mass spectrometry analysis**

Frozen tissue sections (thickness: 15 µm) were prepared from each sample and subjected to DESI MSI. After measurement, tissue sections were H&E stained and underwent histological examination. For DESI analysis, tissue sections were brought to room temperature. DESI MSI was carried out using a home-built DESI ion source coupled to a high resolution orbital trapping mass spectrometer (Exactive, Thermo Scientific GmbH, Bremen, Germany). The DESI ion source was operated at a spatial resolution of 100 µm using nitrogen pressure of 4 bar, high voltage of 4.5 kV, a solvent mixture of methanol and water in a ratio 95:5 and a flow rate of 1.5 µL/min. The mass analysis was carried out in negative ion mode in the mass range m/z = 200-1000 at a nominal mass resolution of 100,000 @ m/z = 200 and a mass accuracy of ± 2 ppm.

**Data analysis**

Raw mass spectrometric data was converted to imzML format (24) via imzML Converter (version 1.0.5) and imported into MATLAB (R2013a) for pre-processing (described in the
supplemental material) and analysis. Pre-processed data was subjected to multivariate statistical data analysis. Recursive Maximum Margin Criterion (RMMC) analysis was used for supervised discrimination and classification (25, 26). Tissue types in each sample and their spatial distribution were determined by histological examination by an independent consultant histopathologist. This information was then used to select a small number of representative mass spectra per tissue type from the MSI image to build a sample specific RMMC model and to classify all pixels of the MSI image via this model. 100 of the correctly classified mass spectra were randomly selected and used to calculate an averaged mass spectrum for the tissue types present in the sample. If several tumor bed samples were measured per patient, average spectra for the tissue types present in the samples were calculated. After this step each patient was represented by a set of average mass spectra corresponding to each tissue type found in the samples taken from this patient. This reduced data set was then subjected to RMMC analysis to answer global questions regarding the differentiation of tissue types, tumor and non-tumor tissue, cancer type, cancer grade and receptor status. Numbers of patients and samples per group of the reported models are given in Supplemental table 1. The histopathological data was retrieved from patient records and was used for data grouping in the corresponding models. Discriminating models were validated using 10-fold cross-validation (described in supplemental material). Confusion matrices were used to show the accuracy of a classification result. In case of two classes, the receiver operating characteristic (ROC) curve and corresponding area under the curve (AUC) were further employed for discriminative model diagnostics. The contributions of m/z features into multivariate models (component weights) were used for ranking of discriminant related features. Additionally, the univariate ANOVA analysis was employed to identify discriminatory m/z variables. The derived p-values were adjusted with the false discovery rate of 0.003 for the tumor/tumor bed model, 0.005 for the tumor/glandular tissue model and 0.01 for the
stroma/tumor stroma and grade/hormone receptor model. After identification of class separating m/z bins via contributions plots and based on p-values, exact m/z values were retrieved from the raw data. These values and isotopic patterns were used for compound identification via data base search (Metlin (27), Lipid maps (28), HMDB (29)).

Results

Cancer diagnostics

We first determined the ability of DESI MSI to identify the presence of breast cancer based on its global metabolic profile from 126 tissue biopsy samples derived from 50 patients. Data extracted from DESI MSI measurements were separated in two groups (tumor from 28 patients and tumor bed from 22 patients) and subjected to multivariate statistical analysis. Tumor bed samples were used as the normal control and were confirmed as morphologically normal tissue by histological examination. Recursive Maximum Margin Criterion (RMMC) analysis demonstrated significant class separation in the cross-validated score plot (Figure 1a). The predictive power of the model was good, with 98.2% of data points correctly classified as tumor and tumor bed (Figure 1b). One tumor sample was incorrectly classified as histologically normal tissue by the model (highlighted by an additional circle in green in Figure 1a). Figure 1c shows the raw summary mass spectrometric profiles of the two groups (tumor and tumor bed). Both tissue profiles feature a high number of peaks (3500 were used to build the model), especially in the mass range of fatty acids (m/z 200-400) and phospholipids (m/z 650-950). Although the majority of features were found in both tissue profiles, the abundance of the individual features differed for the two groups. The contribution plot shown in Figure 1d indicates that fatty acids in general and phospholipids with a mass greater than 760 Da are more abundant in tumor tissue, while a small number of features in the mass range 700-760 Da are significantly higher in the morphologically normal breast tissue. Free fatty acid and
phospholipid markers which have a high contribution to the class separation are summarized in Table 2 and will be described in detail in the discussion section.

**Tissue specific metabolomic profiling**

The main advantage of DESI MSI is that it allows spatially resolved profiling of tissue sections. This requires sophisticated statistical analysis, but allows biochemical changes, for example due to cancer initiation and histological variation, to be visualized and objectively analyzed. Breast tissue is microscopically composed of adipose tissue (approximately 60%), glandular tissue (approximately 30%) and stroma. These were the main tissue components identified during histological evaluation of samples in this study. Accordingly, mass spectrometric profiles for each of these tissue subclasses were extracted per sample and patient. Within our study we examined conventional ‘breast cancer’ – i.e. cases of malignancy in which the neoplastic cells demonstrates an epithelial (glandular) phenotype.

Tumors, in the macroscopic sense, contain multiple tissue types in addition to tumor cells; often in the form of stroma and rarely as adipose tissue. For the purpose of analysis histologically normal stroma and adipose tissue directly associated with malignant cells (macroscopically intratumoral) was extracted separately from this same tissue type located within the tumor bed. Some samples featured blood vessels and nerves with diameters > 100 µm. Due to the low number of samples these were excluded from statistical data analysis. Figure 2a shows the cross-validated RMMC scores plot for the tissue type specific analysis of the data set, differentiating six tissue subclasses: adipose tissue from tumor bed, stroma from tumor bed, glandular tissue from tumor bed, tumor-associated stroma, tumor-associated adipose tissue and tumor itself (neoplastic glandular cells). There was 83.8% overall correct classification (Figure 2b), and despite significant metabolic homogeneity in these tissue type subclasses only a single tumor sample (the
same as in the previous model) was incorrectly classified. This demonstrates that tumor tissue has a distinctly different metabolomic profile in comparison to benign tissues. Of particular note is that tumor demonstrates a different profile to benign glandular tissue. Interestingly tumor-associated stroma was also found to clearly separate from tumor bed stroma. This highlights the fact that stromal tissue within the tumor microenvironment also undergoes a detectable metabolic change. The most significant mass spectrometric differences were observed between stromal tissue and all other tissue types, most likely due to the low abundance of phospholipids in collagen-rich tissue.

The observed differences in metabolomic tissue profiles were sufficient for the classification of unknown tissue types using multivariate statistical methods. Figure 2c shows the results for biopsies taken from two patients. It compares the H&E stained tissue section (stained after DESI MSI measurement) with a tissue map which was reconstructed after classification of each pixel of the DESI MSI measurements using multivariate statistical analysis. For purposes of classification, a model was built based on the metabolomic tissue profiles extracted from all samples leaving out the data of the sample to be classified. The probability of a pixel to belong to a specific tissue type within the model is visualized by intensity of the color representative of that tissue type. Further details of model generation and classification can be found in the supplemental material. On the left of the Figure, an H&E stained tumor bed biopsy with histologically normal stromal, glandular and adipose tissue and the associated DESI MSI tissue map is shown. On the right, an H&E stained tissue section of grade 3, HR negative, IDC (with necrosis and a large amounts of tumor invading stroma in the superior aspect of the tissue section, and histologically normal stroma and adipose tissue in the inferior aspect). The associated DESI MSI tissue map is adjacent. Tissue types, as identified in the H&E stained section, were also accurately classified with low lateral resolution by multivariate statistical analysis.
of the metabolomic data, thus showing the potential of this method for the detection of tumor tissue without histological staining. The method shown here is proof-of-concept and could be improved by adding more tissue types and samples to the model.

In order to further investigate the previously described peritumoral metabolomic changes associated with breast cancer, glandular tissue was specifically examined by statistical data analysis. This tissue specific comparison allows any direct changes in the metabolome of the glandular tissue due to cancer development to be characterized, without interference by the metabolome of adjacent tissue, as often observed in bulk tissue analysis. DESI MSI as performed here is not a quantitative method, but allows for relative quantification when, as shown below, if data from the same tissue type is compared. In this study mass spectrometric profiles of morphologically normal glandular tissue (18 patients) were compared to those of tumor tissue (28 patients). The contribution of each signal to the separation of the groups was calculated. The vast majority of peaks were found to be more abundant in the tumor tissue. The most intense peaks (with a p-value < 1·10^{-5}) are summarized in Table 2 with tentative identifications along with mean values and fold change for the morphologically normal glandular breast tissue, and tumor tissue. Among the highly significant peaks identified were several fatty acids (FA), phosphatidylethanolamines (PE), phosphatidylcholines (PC) and phosphatidylinositols (PI). Fold changes between 2 and 14 were observed (mean fold change = 5), which is close to fold changes reported in the literature (30, 31). Only a small number of peaks with medium significance for tissue separation were found in greater abundance in the morphologically normal glandular tissue. These signals were identified as the ether lipids p-PE(36:4), p-PE(38:4), p-PE(38:5) and PA(36:1).

In order to elucidate the metabolomic changes associated with cancer progression and invasion as found in DESI MS spectra in more detail, stroma was examined independently
by statistical data analysis. In our study mass spectrometric profiles of morphologically normal stroma from the tumor bed (20 patients) were compared to stroma directly associated with the tumor (20 patients). Again, there was a greater abundance of signals from fatty acids and lipids in ‘tumor associated’ stroma compared to stroma from histologically normal tissue. The most intense peaks, with a p-value < 2·10^{-5}, are summarized in Table 3. Very high fold changes (26-37) were found for signals associated with lactate and calcidiol. Lactate was found as a dimer and in form of several sodium chloride clusters (M+Na₄Cl₄, M+Na₃Cl₃, M+Na₂Cl₂). Salt clusters are likely to be observed in stroma, since this tissue serves as a water and salt depot for the surrounding tissues (32). The monomer however was not accessible in the given measurement range.

**Associations of metabolomic profiles with clinical factors**

Finally, the group of tumor samples was divided into subgroups (see Table 1) according to a range of clinical factors employed in the prognostication and management of breast cancer. Among these were the histological grade and type of tumor, lymph node status, hormone receptor (estrogen (ER) and progesterone (PR)) status and HER2 receptor status. Each of these clinical factors was tested individually (independently of all other clinical factors) to fit the metabolomic differences observed in the data of the tumor tissue. Patients were excluded from the data set when no data for the tested clinical factor was available. Differences in the metabolic data were best explained by tumor grade (Figure 3) and hormone receptor status (Figure 4) and slightly less accurately by tumor type (see Supplemental Figure 1).

The correlation of the metabolomic data with the tumor grade is shown in Figure 3 for tumor grade 3, tumor grade 2 and morphologically normal glandular tissue. The cross-validated RMMC score plot for this data set is shown in Figure 3a. The corresponding confusion matrix is given in Figure 3b. An overall classification accuracy of 88.6% was
achieved. The ROC curve shows the potential to use the metabolomic data for the
discrimination of grade 2 and grade 3 tumors (Figure 3c). The probability for correct
discrimination was 97% (AUC = 0.97).

The correlation of the metabonomic data with hormone receptor status (HR) of the tumor
tissue is shown in Figure 4 for tumors with positive hormone receptor status, tumors with
negative hormone receptor status and morphologically normal glandular tissue. An overall
classification accuracy of 86.7% was achieved for this model (Figure 4a and b). The
probability for correct discrimination of HR positive and HR negative tumors was 96%
(AUC = 0.96, Figure 4c).

Tumor grade und hormone receptor status fitted the metabolomic data with similar quality
due to the 86% overlap of the two clinical factors. Hence a combination of tumor grade and
hormone receptor was used for further interpretation.

Supplemental Figure 2 shows intensity distribution plots of 8 compounds with p-values
< 2·10^{-4}. Supplemental Figure 3 shows the distribution of FA(20:2) in a tumor bed sample
(a), in a grade 2 HR positive tumor sample (b) and a grade 3 HR negative tumor sample
(c). From a clinical point of view a separation between grade 2, HR positive tumors and
grade 3, HR negative tumors is relevant, since the former are generally associated with a
better response to adjuvant therapy and a better prognosis, whilst the opposite is true of
the latter (33).

Discussion

Cancer diagnostics

Results correspond well with previous measurements of the metabolic profile of breast
cancer. Studies using both using mass spectrometry (30) and NMR spectroscopy (34)
have documented an increase in both fatty acid and phospholipid concentrations in breast
cancer tissue in comparison to morphologically normal breast tissue. The single misclassified sample in our study showed lower overall intensity in the average spectrum when compared to the other tumor samples. Lower overall intensity might be related to the low density of malignant cells in the tissue section. No clinical (tumor type, tumor grade, ER, PR, HER2, lymph node status, preoperative treatment) or demographic (age, ethnicity, lifestyle) features were found which explain the misclassification.

**Alterations in stroma and glandular tissue due to cancerization and immune response**

Cancer is generally associated with a loss of differentiation, unregulated growth and increased cell mobility. Research into the mechanisms involved in the transformation of normal epithelial cells into cancer cells has already determined dramatic changes in the genome(35, 36), proteome(37, 38) and metabolome (39, 40) of these cells. Among these changes, an alteration in the glucose metabolism (41), in the fatty acid (31, 34) as well as in the lipid composition of the cells(30) has been observed. In particular, fatty acids from de novo synthesis were found in greater abundance. These findings were replicated in our data set where five of the seven fatty acids found to be increased in tumor tissue can be generated via de novo synthesis. Higher PC and PE levels have been reported previously in breast cancer (31, 42, 43) and these could be associated with the loss of differentiation and, more specifically, a loss of epithelial polarity. Previous studies have shown that PEs and PCs were up-regulated, while p-PEs were down-regulated in Madin–Darby canine kidney cells (a cell culture model for epithelial polarization) after the cells underwent epithelial-mesenchymal transition (42).

There are similarities in the immune response towards wounds and towards tumors, such as the activation of extracellular matrix remodeling (EMC), cell motility and angiogenesis (44). Furthermore, there is increasing evidence that the tumor environment plays an active
role in tumor development and progression (45). Tumor cells are known to interact with their surrounding tissue and may manipulate the function of cells such as fibroblasts, endothelial cells and pericytes (46, 47). Fibroblasts are highly abundant in the tumor environment (47). They have been shown to undergo differentiation to cancer-associated fibroblasts (CAFs) and to produce lactate and pyruvate through aerobic glycolysis - prompting tumor cells to overcome energy depletion (reverse Warburg effect) (48). But tumor cells themselves are also known to provide for their increased energy consumption by producing lactate (Warburg effect). The increased levels of lactate found within the tumor stroma in our study can thus be explained by invading cancer cells or CAFs. In addition we found in our study an up-regulation of fatty acids and phospholipids in the intratumoral stroma. This may be explained by the increased cellularity of the tumor associated stroma due to immune response and inflammation, as well as malignant cell invasion. Fatty acids and lipids showed similar fold changes to those observed in the tumor tissue itself, which may be an indication that these signals are mainly from invading tumor cells.

**Associations of metabolomic profiles with clinical factors**

Our findings were consistent with a HPLC-ESI-MS study of 257 breast cancer samples and 10 control samples (30). Seven of the eleven lipids which we found to be up-regulated in cancer tissue were also found by the HPLC-ESI-MS study. Furthermore, these lipids showed the same association to tumor grade and hormone receptor status in terms of relative compound concentrations in our data as was described in the HPLC-ESI-MS study.

**In conclusion**, using DESI MSI we were able to differentiate between tumor tissue and morphologically normal breast tissue with an overall accuracy of 98.2%. In comparison to morphologically normal glandular tissue, tumor tissue had increased levels of fatty acids.
and showed an increase in the concentration of the majority of measured phospholipids. These findings support the theory of de novo lipogenesis in tumor tissue and were associated with loss of differentiation and increased levels of tumor cell proliferation. In addition, we demonstrated changes in tumor-associated stroma which correlate with findings that tumor supporting stroma may be associated with disease prognosis by playing a role in angiogenesis and extracellular matrix remodeling. We found that tumor grade and hormone receptor status accounted for the greatest differences in the metabolic profiles of tumor tissue, with tumor type being less significant. The combination of tumor grade and hormone receptor status allowed for the differentiation of normal glandular tissue, grade 2 tumors with a positive hormone receptor status and grade 3 tumors with negative hormone receptor status with an overall accuracy of 87.7%. Therefore DESI MSI may prove to be a useful tool in the future diagnosis of breast cancer and spatially resolved characterization of tumor associated changes in lipid biochemistry.

Acknowledgements

The work was funded by the European Research Council under the Starting Grant scheme (contract no. 210356) and the European Commission FP7 Intelligent Surgical Device project (contract no. 3054940). The authors would like to acknowledge the NIHR Biomedical Research Centre for funding surgical metabonomics projects at Imperial College London. KV acknowledges a junior research fellowship scheme for supporting this research project.

References


Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.


### Table 1

<table>
<thead>
<tr>
<th></th>
<th>N patients</th>
<th>(samples)</th>
<th>%</th>
<th>Median age (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tumor bed biopsies</strong></td>
<td>22</td>
<td>(98)</td>
<td>44</td>
<td>55 (17-87)</td>
</tr>
<tr>
<td><strong>Tumor biopsies</strong></td>
<td>28</td>
<td>(28)</td>
<td>56</td>
<td>61 (44-90)</td>
</tr>
<tr>
<td><strong>Tumor type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invasive ductal carcinoma (IDC)</td>
<td>19</td>
<td></td>
<td>68</td>
<td>63 (44-90)</td>
</tr>
<tr>
<td>Invasive lobular carcinoma (ILC)</td>
<td>9</td>
<td></td>
<td>32</td>
<td>59 (48-86)</td>
</tr>
<tr>
<td><strong>Tumor grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>2</td>
<td></td>
<td>7</td>
<td>61 (53-69)</td>
</tr>
<tr>
<td>T2</td>
<td>13</td>
<td></td>
<td>46</td>
<td>62 (44-87)</td>
</tr>
<tr>
<td>T3</td>
<td>13</td>
<td></td>
<td>46</td>
<td>59 (45-90)</td>
</tr>
<tr>
<td><strong>Hormone receptor (HR) status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td>12</td>
<td></td>
<td>43</td>
<td>61 (45-90)</td>
</tr>
<tr>
<td>positive</td>
<td>15</td>
<td></td>
<td>54</td>
<td>59 (44-87)</td>
</tr>
<tr>
<td><strong>HER2 receptor status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td>20</td>
<td></td>
<td>71</td>
<td>57 (44-87)</td>
</tr>
<tr>
<td>positive</td>
<td>7</td>
<td></td>
<td>25</td>
<td>64 (51-90)</td>
</tr>
<tr>
<td><strong>Lymph node status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td>13</td>
<td></td>
<td>29</td>
<td>59 (44-85)</td>
</tr>
<tr>
<td>positive</td>
<td>13</td>
<td></td>
<td>61</td>
<td>62 (45-90)</td>
</tr>
<tr>
<td>compound</td>
<td>ion</td>
<td>m/z value</td>
<td>mean intensity</td>
<td>mean intensity</td>
</tr>
<tr>
<td>----------</td>
<td>-----</td>
<td>-----------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>FA(16:1)</td>
<td>M-H</td>
<td>253.22</td>
<td>149</td>
<td>37</td>
</tr>
<tr>
<td>FA(18:1)</td>
<td>M-H</td>
<td>281.25</td>
<td>2361</td>
<td>648</td>
</tr>
<tr>
<td>FA(20:3)</td>
<td>M-H</td>
<td>305.25</td>
<td>281</td>
<td>41</td>
</tr>
<tr>
<td>FA(20:2)</td>
<td>M-H</td>
<td>307.27</td>
<td>131</td>
<td>13</td>
</tr>
<tr>
<td>FA(20:1)</td>
<td>M-H</td>
<td>309.29</td>
<td>181</td>
<td>22</td>
</tr>
<tr>
<td>FA(22:5)</td>
<td>M-H</td>
<td>329.25</td>
<td>114</td>
<td>37</td>
</tr>
<tr>
<td>FA(22:4)</td>
<td>M-H</td>
<td>331.27</td>
<td>149</td>
<td>36</td>
</tr>
<tr>
<td>FA(24:1)</td>
<td>M-H</td>
<td>365.35</td>
<td>57</td>
<td>4</td>
</tr>
<tr>
<td>LPI(18:1)</td>
<td>M-H</td>
<td>599.33</td>
<td>79</td>
<td>13</td>
</tr>
<tr>
<td>PE(32:0)</td>
<td>M-H</td>
<td>690.51</td>
<td>52</td>
<td>11</td>
</tr>
<tr>
<td>PE(35:1)</td>
<td>M-H</td>
<td>730.55</td>
<td>79</td>
<td>13</td>
</tr>
<tr>
<td>PC(34:2)</td>
<td>M+Cl</td>
<td>792.54</td>
<td>264</td>
<td>125</td>
</tr>
<tr>
<td>PC(34:1)</td>
<td>M+Cl</td>
<td>794.55</td>
<td>519</td>
<td>166</td>
</tr>
<tr>
<td>PS(38:3)</td>
<td>M-H</td>
<td>812.55</td>
<td>428</td>
<td>199</td>
</tr>
<tr>
<td>PC(36:1)</td>
<td>M+Cl</td>
<td>822.59</td>
<td>122</td>
<td>33</td>
</tr>
<tr>
<td>PI(36:3)</td>
<td>M-H</td>
<td>859.54</td>
<td>377</td>
<td>127</td>
</tr>
<tr>
<td>PI(38:3)</td>
<td>M-H</td>
<td>887.57</td>
<td>1213</td>
<td>370</td>
</tr>
<tr>
<td>PI(40:5)</td>
<td>M-H</td>
<td>911.57</td>
<td>177</td>
<td>76</td>
</tr>
<tr>
<td>PI(40:4)</td>
<td>M-H</td>
<td>913.59</td>
<td>182</td>
<td>53</td>
</tr>
<tr>
<td>compound</td>
<td>ion</td>
<td>m/z value</td>
<td>mean intensity</td>
<td>mean intensity</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------</td>
<td>-----------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>lactate</td>
<td>2M+Na</td>
<td>201.04</td>
<td>84</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>M-H</td>
<td>305.25</td>
<td>82</td>
<td>12</td>
</tr>
<tr>
<td>lactate</td>
<td>M+Na₄Cl₂</td>
<td>320.86</td>
<td>35</td>
<td>1</td>
</tr>
<tr>
<td>Calcidiol</td>
<td>M-2H+Na</td>
<td>421.32</td>
<td>36</td>
<td>1</td>
</tr>
<tr>
<td>LPS(20:3)</td>
<td>M-H₂O-H</td>
<td>528.28</td>
<td>43</td>
<td>8</td>
</tr>
<tr>
<td>PC(38:4)</td>
<td>M+Cl</td>
<td>844.57</td>
<td>78</td>
<td>14</td>
</tr>
<tr>
<td>PI(36:1)</td>
<td>M-H</td>
<td>863.57</td>
<td>129</td>
<td>38</td>
</tr>
</tbody>
</table>
**Table and Figure Legends**

**Table 1.** Demographic and Clinicopathological Characteristics

* Information about HR, HER2 receptor and lymph node status was not available for one patient. Lymph node status was also not available for another patient.

**Table 2:** Compounds found most discriminating between morphologically normal glandular tissue and tumor tissue (FDR: 0.005)

**Table 3:** Compounds most significant in discriminating between morphologically normal stroma of the tumor bed and tumor-associated stroma (FDR: 0.01)

**Figure 1:** Differentiation of tumor (red) and tumor bed samples (green) based on metabolomic profiles measured with DESI MSI. a) 10-fold cross-validated RMMC score plot (each patient is max. represented by one point in each group, for each point max. 100 randomly selected spectra were averaged) showing the separation of the two groups using multivariate analysis, b) Confusion matrix showing classification results based on mahalanobis distance (98.2% overall correct assignment), c) Representation of the tissue specific metabolomic profiles, d) Contribution plot showing the contribution of each signal in the metabolomic profiles to the group separation by a color bar (lower p-values indicate higher significance for group separation). False discovery rate was 0.003 or <1 sample for signals in red color.

**Figure 2:** Differentiation of tissues present in tumor and tumor bed samples based on the metabolomic profiles measured with DESI MSI. a) 10-fold cross-validated RMMC score plot (each patient is max. represented by one point in each group, for each point max. 100 randomly selected spectra were averaged) showing the separation of the groups using multivariate analysis, b) Confusion matrix showing classification results based on mahalanobis distance (83.8% overall correct assignment) c)
Comparison of H&E stained tissue section (stained after DESI MSI) used for histological evaluation and tissue classification results from multivariate statistical analysis of the DESI MSI metabolomic data. Colored pixels correspond to tissues as indicated in the Figure legend. Left panel shows a tumor bed biopsy. Right panel shows a grade 3 IDC with negative hormone receptor status.

**Figure 3:** Association of metabolomic DESI MSI profiles of normal glandular and tumor (ILC and IDC) tissue with tumor grade. a) 10-fold cross-validated RMMC score plot (each patient is max. represented by one point in each group, for each point max. 100 randomly selected spectra were averaged) showing the separation of normal glandular tissue, grade 2 tumors and grade 3 tumors using multivariate data analysis, b) Confusion matrix of the RMMC model showing classification results based on mahalanobis distance (88.6% overall correct assignment), c) ROC curve showing the potential of the model to discriminate between grade 2 tumors and grade 3 tumors.

**Figure 4:** Association of metabolomic DESI MSI profiles of normal glandular and tumor tissue with hormone receptor status. a) 10-fold cross-validated RMMC score plot (each patient is max. represented by one point in each group, for each point max. 100 randomly selected spectra were averaged) showing the separation of normal glandular tissue, tumors with positive HR status (ER/PR receptor positive) and tumors with negative HR status (ER/PR receptor negative) using multivariate data analysis, b) Confusion matrix showing classification results based on mahalanobis distance (86.7% overall correct assignment), c) ROC curve showing the potential of the model to discriminate between positive and negative HR status.
Figure 1

a) RMMC

b) Confusion matrix

- Predicted Class
  - 96.4%, 27
  - 3.6%, 1
  - 0%, 0
  - 100%, 22

Target Class

- Tumor
- Tumor bed

Tissue specific mass spectra

m/z

Covariance(X,y)

m/z

FDR = 0.003, anova

p-values

Tumor

Tumor bed
Figure 2

a) RMMC

b) Confusion matrix

<table>
<thead>
<tr>
<th>Target Class</th>
<th>Adipose tissue from tumor bed</th>
<th>Tumor-associated adipose tissue</th>
<th>Stroma from tumor bed</th>
<th>Tumor stroma</th>
<th>Glandular tissue from tumor bed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81.8%, 18</td>
<td>0%, 0</td>
<td>18.2%, 4</td>
<td>0%, 0</td>
<td>0%, 0</td>
<td>0%, 0</td>
</tr>
<tr>
<td>0%, 0</td>
<td>85.7%, 6</td>
<td>0%, 0</td>
<td>0%, 0</td>
<td>0%, 0</td>
<td>14.3%, 1</td>
</tr>
<tr>
<td>9.5%, 2</td>
<td>0%, 0</td>
<td>81.0%, 17</td>
<td>4.8%, 1</td>
<td>4.8%, 1</td>
<td>0%, 0</td>
</tr>
<tr>
<td>0%, 0</td>
<td>4.8%, 1</td>
<td>4.8%, 1</td>
<td>71.4%, 15</td>
<td>0%, 0</td>
<td>19.0%, 4</td>
</tr>
<tr>
<td>0%, 0</td>
<td>0%, 0</td>
<td>11.1%, 2</td>
<td>0%, 0</td>
<td>83.3%, 15</td>
<td>5.6%, 1</td>
</tr>
<tr>
<td>0%, 0</td>
<td>0%, 0</td>
<td>0%, 0</td>
<td>0%, 0</td>
<td>3.6%, 1</td>
<td>96.4%, 27</td>
</tr>
</tbody>
</table>

C) Classification of unknown breast tissue from RMMC model

H&E stained section:  
- Adipose Tissue  
- Glandular Tissue  
- Stroma

Tissue map:  
- Adipose tissue (tumor bed)  
- Glandular tissue (tumor bed)  
- Tumor stroma  
- Tumor-associated adipose tissue  
- Tumor

1mm
Figure 3

a) RMMC

![RMMC Diagram](image)

- Red: Glandular tissue from tumor bed
- Green: Tumor, grade 2
- Blue: Tumor, grade 3

b) Confusion matrix

<table>
<thead>
<tr>
<th>Predicted Class</th>
<th>Target Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glandular tissue from tumor bed</td>
<td>88.9%, 16</td>
</tr>
<tr>
<td>Tumor, grade 2</td>
<td>7.7%, 1</td>
</tr>
<tr>
<td>Tumor, grade 3</td>
<td>0.0%, 0</td>
</tr>
</tbody>
</table>

c) ROC curve

![ROC Curve](image)

AUC = 0.97
Figure 4

(a) RMMC

![Graph showing components with Glandular tissue from tumor bed, Tumor, HR positive, and Tumor, HR negative.

(b) Confusion matrix

<table>
<thead>
<tr>
<th>Predicted Class</th>
<th>Target Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glandular tissue from tumor bed</td>
<td>88.9%, 16</td>
</tr>
<tr>
<td>Tumor, HR positive</td>
<td>0.0%, 0</td>
</tr>
<tr>
<td>Tumor, HR negative</td>
<td>6.7%, 1</td>
</tr>
</tbody>
</table>

(c) ROC curve

AUC = 0.96
Spatially resolved metabolic phenotyping of breast cancer by desorption electrospray ionization mass spectrometry

Sabine Guenther, Laura J Muirhead, Abigail VM Speller, et al.

Cancer Res  Published OnlineFirst February 17, 2015.

Updated version  Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-14-2258

Supplementary Material  Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2015/02/17/0008-5472.CAN-14-2258.DC1

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.