Imaging of esophageal lymph node metastases by desorption electrospray ionization mass spectrometry

Nima Abbassi-Ghadi¹, Ottmar Golf², Sacheen Kumar¹, Stefan Antonowicz¹, James S McKenzie², Juzheng Huang¹, Nicole Strittmatter², Hiromi Kudo³, Emrys A Jones², Kirill Veselkov², Robert Goldin³, Zoltan Takats², George B Hanna¹

1 Department of Surgery and Cancer, Imperial College London, 10th Floor QEQM Wing, St Mary’s Hospital, London, UK, W2 1NY
2 Department of Surgery and Cancer, Imperial College London, 6th Floor Sir Alexander Fleming Building, South Kensington Campus, London, UK, SW7 2AZ
3 Centre for Pathology, Imperial College London, 4th Floor Clarence Wing, St Mary’s Hospital, London, UK, W2 1NY

Running Title: DESI-MSI of esophageal lymph node metastases

Keywords: DESI-MSI, lymph node metastases, esophageal adenocarcinoma, glycerophospholipids, lipidomics.

Financial support: The authors would like to acknowledge funding from the European Research Council (DESI-JeDI Imaging Starting Grant; MASSLIP Consolidator Grant), the National Institute of Health Research (Imperial Biomedical Research Centre) and the National Institute of Health Research - Diagnostic Evidence Cooperatives London.

Corresponding Author:
Professor Zoltan Takats
Imperial College London
371 Sir Alexander Fleming Building
South Kensington Campus, London, UK, SW7 2AZ
Tel/Fax: (+) 44 2075942760
z.takats@imperial.ac.uk

Conflicts of interest: No conflicts of interest are disclosed

Word Count: 4194  Total number of figures and tables: 6 + 1
**Abstract**

Histopathological assessment of lymph node metastases (LNM) depends on subjective analysis of cellular morphology with inter-/intra-observer variability. In this study, LNM from esophageal adenocarcinoma was objectively detected using desorption electrospray ionization-mass spectrometry imaging (DESI-MSI). Ninety lymph nodes and their primary tumor biopsies from 11 esophago-gastrectomy specimens were examined and analyzed by DESI-MSI. Images from mass spectrometry and corresponding histology were co-registered and analyzed using multivariate statistical tools. The MSIs revealed consistent lipidomic profiles of individual tissue types found within lymph nodes. Spatial mapping of the profiles showed identical distribution patterns as per the tissue types in matched immunohistochemistry images. Lipidomic profile comparisons of LNM versus the primary tumor revealed a close association in contrast to benign lymph node tissue types. This similarity was used for the objective prediction of LNM in mass spectrometry images utilizing the average lipidomic profile of esophageal adenocarcinoma. The multivariate statistical algorithm developed for LNM identification demonstrated a sensitivity, specificity, positive predictive value and negative predictive value of 89.5, 100, 100 and 97.2 per-cent, respectively, when compared to gold-standard immunohistochemistry. DESI-MSI has the potential to be a diagnostic tool for peri-operative identification of LNM and compares favorably with techniques currently used by histopathology experts.
**Introduction**

Lymphatic spread is the most common route of cancer dissemination (1, 2). The presence of metastases in regional lymph nodes (LN) indicates cancer progression and is an important prognostic factor for long-term survival (3-6). Evaluation of lymph node metastases (LNM) forms an integral component of cancer staging systems (7, 8), which are routinely used for decision-making in treatment.

Curative therapy for the majority of carcinomas consists of complete surgical resection of the primary tumor and an appropriate regional LN clearance. The sentinel LN is considered the gatekeeper to regional LNs, and is likely to be the first to harbor metastases (9-11). The concept of the sentinel LN has been used intra-operatively to guide the extent of LN clearance in several tumor sites, including breast, thyroid, melanoma, esophageal and gastric cancer (12-15). It is identified at time of surgery and assessed for the presence of tumor; the histopathological evaluation of which guides the surgeon in deciding whether to proceed with clearance of the rest of the regional LNs.

Intra-operative histopathological analyses of LNs using frozen section or touch imprint cytology have sensitivities of 73 and 63 per-cent and specificities of 100 and 99 per-cent for the identification of LNM, respectively, when compared to the gold standard paraffin embedded tissue based analysis as determined by meta-analyses (16, 17). Histopathological assessment relies on subjective analysis of cellular morphology resulting in inter- and intra-observer variability (18, 19). Although immuno-histochemistry of paraffin embedded tissue has been shown to improve sensitivity for occult metastases in histopathological analysis, it has a lengthier processing time and does not suit intra-operative assessment. The presence of LNM in paraffin sections of sentinel LNs, wrongly diagnosed as negative for cancer during intra-operative assessment, necessitates patient to have another surgery to remove regional
LNs resulting in patient distress, possible complications and cost implications for healthcare providers.

Several non-histopathological techniques have been used to assess LN status intraoperatively. The use of one-step nucleic acid amplification (OSNA) assay (20, 21), real-time quantitative reverse-transcription polymerase chain reaction (22), and technologies such as photo-acoustic tomographic imaging (23), hand held positron emission tomography probe (24) and Raman spectroscopy (25) have been reported in this context. The only method deemed suitable for clinical application is the OSNA assay in breast cancer, which relies on the molecular identification of cytokeratin-19 (CK19) mRNA. However, when processing a whole LN using the OSNA assay, several important features including size, location and pattern of LNM as well as extra-capsular extension are lost. In addition, no residual LN tissue remains after OSNA assay should further analysis be required.

Novel methods of histological classification, based on the chemical organization of tissue samples, using mass spectrometry imaging (MSI), have gained increasing interest. Tumor versus normal surrounding tissue, have been demonstrated with false color images representative of specific chemical profiles (26, 27) and in close correlation to the matched histopathological image. Commonly used MSI techniques include matrix-assisted laser desorption/ionization (MALDI), secondary ion mass spectrometry (SIMS) and DESI-MSI (28, 29). DESI-MSI is an ideal technology for the development of targeted objective histology workflows, such as LNM identification, as it requires minimal sample preparation as direct tissue analysis can be performed under ambient conditions (30). Also, DESI-MSI is a non-destructive technique allowing further histopathological analysis if needed.

This study reports the application of DESI-MSI for the identification of esophageal adenocarcinoma LNM. Comparative analysis of the lipidomic profiles of the primary tumor
and metastases versus benign lymph node tissue was carried out. The similarities in lipidomic profiles between primary tumor and LNM was utilized to identify LNM based on an average primary tumor profile using multivariate methods. The proposed workflow for intra-operative analysis would involve the retrieval of the lymph node followed by MSI and an integrated protocol for the objective identification of metastases based on its lipid profile, independent of expert histopathological opinion.
**Materials and Methods**

**Clinical specimens and patient selection**

Approval for the study was obtained from the institutional ethics review committee and Imperial College Healthcare Tissue Bank (Project number R14120). Consecutive patients undergoing transthoracic esophago-gastrectomy for EA were recruited into this study. Exclusion criteria included patients with esophageal squamous cell carcinoma, malignancy associated with any other site in the body, liver disease and patients with signs/symptoms of acute infection.

Manual LN dissection was performed immediately after retrieval of the surgical specimen. The harvested LNs were divided in two along their long axis. For each LN, one half was sent for routine histopathological examination and the other half was snap frozen in liquid nitrogen for MSI. Incision biopsies were also taken from the primary tumor for the same purposes.

**Sample preparation, DESI-MSI and reference test**

Cryo-sectioning and DESI-MSI were performed in line with methods described in our previous study (31).

**Histopathological assessment: reference test**

The same cohort of LNs used for DESI-MSI, were subsequently stained with haematoxylin & eosin (H&E) and also immuno-stained with the anti-cytokeratin AE1/AE3 antibody (Dako Ltd, Ely, UK) to detect the presence of LNM. The samples of primary tumor were only stained with H&E. Digital images were acquired for bioinformatics analysis using a high resolution slide scanner (NanoZoomer2.0-HT, C9600-13 Hamamatsu, Hamamatsu City, Japan). Two histopathologists, specializing in esophago-gastric cancer, blinded to the results
of the DESI-MSI, assessed the LN sections for the presence of macro-metastases (>2mm) and micro-metastases (0.2mm-<2mm) (32, 33) using standard bright field microscopy. The identification of LNM was aided by the red/brown staining of the AE1/AE3 anti-cytokeratin stain. Any disagreement in the results were resolved by a third histopathologist. The location of the metastases were mapped onto the digital images.

**Data Analysis**

Refer to supplementary materials and methods for full methodological details of data analysis with respect to: tissue specific mass spectra extraction; data pre-processing (mass range selection, peak alignment, normalization, de-noising, data averaging); multivariate statistical models; glycerophospholipid (GPL) annotation and individual GPL comparison between tissue types. Tissue specific mass spectra extraction was performed from the full dataset, for the following tissue types: LN parenchyma (LNP), carbon deposits, LN connective tissue (LNC), fat, micro-metastases (0.2-2mm), macro-metastases (>2mm), metastases with response to chemotherapy (LNMR) and primary tumor, for the purpose of multivariable analysis.

**Data processing for spatial prediction of metastases**

For the training set, all mass spectra of the non-malignant LN tissue types (LNP, LNC, carbon deposits, fat) were combined into one class, the primary tumor (excluding metastases) into a second class and the glass slide background as a third class.

The full dataset consisted of 11 samples of primary tumor (one from each patient) and the respective LNs from each tumor specimen. A variable number of LNS were retrieved for each primary tumor resulting in a total of 90 LNs. Leave-one-patient-out cross validation was performed in the following manner. A training set was created by excluding a primary tumor and its respective LNs from one of the 11 sample sets, while the LN samples of the excluded
patient represented the test set. Therefore, data from the same patient was not present in the training and test sets at the same time. This had been carried out iteratively for each of the patients to complete the leave-one-patient-out cross validation process.

The m/z values of the training set classes were combined to a common m/z-vector by an in-house peak-matching algorithm. The m/z-values of the test sets were then matched to those of the common m/z-vector of the training set to enable multivariate comparison between the two sets. Each mass spectrum was normalized to its total ion current by dividing each peak intensity of a mass spectrum by the sum of all peak intensities of the same spectrum. M/z values with zero median intensities for all histological classes were considered noise and thus excluded. Variance stabilizing normalization was carried out by log transformation of the data (34).

Multivariate classification of the test set pixels was based on a combination of linear discriminant analysis with recursive maximum margin criterion previously described in detail (35). The primary aim was to classify each pixel (mass spectrum) of the test set as either healthy, tumor or glass slide. Since the test set samples do not contain primary tumor, but only metastases from the primary tumor, the classification of the metastases mass spectra of the test set were expected to succeed based on molecular ion patterns of the primary tumor mass spectra.

Multivariate comparison of the two sets was carried out by projecting the test set mass spectra into the multivariate space of the training set model. This was achieved by multiplication of the test set mass spectra with the training model weights to obtain the score values of the test set. The test set score values were then translated to probabilities by means of multinomial logistic regression. If the highest relative class-membership probability of a mass spectrum (pixel) was below 70% (The sum of all class probabilities are 100 %.), then
the pixel was considered an outlier; otherwise it was labelled as the class with the highest probability. Since the training set only contained the three classes (i.e., primary tumor, healthy tissue and glass slide) the pixels of the test samples were color-coded based on their classification result into either cyan (healthy), red (metastasis based on primary tumor molecular ion patterns), white (glass slide) or grey (outlier). The color-coding of each pixel on a test set creates a tissue class prediction image (TCPI), which allows a spatial location/tissue type accuracy comparison with the matched IHC images of the reference test.

**Comparison of outcomes between the IHC reference test and MSI Tissue classification prediction images**

In order to compare the MSI-based tissue classification method to the IHC reference test, the threshold for a positive diagnosis of metastases on the pixel-wise tissue class prediction images had to be determined. Since the size definition of micro-metastases is 0.2-2mm in its largest dimension, an equivalent value was sought with respect to the imaging resolution of the pixels. In this dataset, each pixel represented 75um in its lateral dimension and therefore a minimum of three adjacent red pixels was defined as a criterion for the positive identification of micro-metastases. Anything less than this pre-analysis threshold was considered normal lymph node tissue. This analysis was carried out by two independent assessors and disagreements were resolved by a third assessor. The binary outcome of the objective histology (metastases versus no metastases) was compared to the binary outcome of the IHC reference test (metastases versus no metastases) by means of a contingency table. Any lymph nodes identified with Isolated Tumor Cells (ITCs, <0.2mm, n=1) on IHC were classified as normal LNds due to poor identification accuracies and lack of clinical significance of this cell type.
Results

A total of 90 lymph nodes were harvested from 11 patients with EA. All lymph nodes and primary tumor samples were included in the analysis with no exclusions. The demographics and case characteristics of the 11 patients and their respective lymph node harvest are shown in table 1.

The 90 LNs comprised of 65 normal LNs (15 containing carbon deposits from smoking), 5 LNs with micro-metastases, 13 LNs with macro-metastases, one LN with ITCs, 4 LNMR and 1 LN with evidence of both macro-metastases and LNMR. Successful DESI-MSI was performed on all LNs and the 11 primary tumor samples. The LN sections subjected to DESI-MSI, were then successfully stained with AE1/AE3 anti-cytokeratin antibody for the purpose of comparative analysis.

Intra-sample comparison of LN tissue types

Figures 1 (a-g) & 2 (a-g) show the intra-sample comparison of the lipidomic profile of specific tissue types found within individual LNs. Figure 1 shows a single LN with LNP, LNC, macro-metastases and LNMR tissue types. The mass spectral profiles (600-1000 m/z) of the tissue classes are distinct and demonstrated in the RMMC scores plot (refer to Supplementary materials and methods for full methodological details). The Principal Component (PC) 1 scores image differentiates the benign LNP and LNC from the malignant macro-metastases and LNMR. The PC 2 scores image further differentiates the LNMR from the other LN tissue types in the same spatial orientation as per the IHC image. Figure 2, shows a lymph node with micro-metastases, which is differentiated from the benign LN tissue types on the PC 3 scores image. This demonstrates the resolution limits of the instrument in detecting mass spectral changes consistent with micro-metastases. Further
examples of individual LN analysis including a LN with carbon deposits and a LN with macro-metastases is provided in Supplementary Figures S1 and S2, respectively.

**Inter-sample comparison of LN tissue types**

Inter-sample comparison was performed to interrogate the lipidomic profile of EA primary tumor (n=11) and its respective LNM (n=19). PCA did not demonstrate distinct clustering of the two tissue types, suggesting that the mass spectral profiles are indistinct (Figure 3a/b). Of the 203 Glycerophospholipids (GPLs) identified in the 600-1000 m/z range, only two were found to be significantly different between the LNM and the primary tumor (Supplementary Table 1). The mean relative abundance±SD of PG (40:6) was 31.6±24.8 and 63.2±24.0 in primary tumor and metastases, respectively (log₂ mean Fc, 1.00; q=0.0265). The mean relative abundance±SD of PG (38:6) was 23.1±21.7 and 69.2±42.1 in primary tumor and metastases, respectively (log₂ mean Fc, 1.58; q=0.0306). Inter-sample comparison of LNM (n=19) versus metastases with LNMR (n=5) is shown in Figure 3c/d. PCA demonstrated separation of the tissue classes and leave one out internal cross validation, with Mahalanobis distance classifier, demonstrated a classification accuracy of 100% for metastases and 80% for LNMR (Supplementary Figure S3). The overall data suggest that primary tumor and LNM share similar lipidomic profiles and that LNMR are distinct from LNM.

Principal components analysis comparing lipidomic profiles (m/z 600-1000) of specific non-malignant LN tissue types (LNP, LNC, fat, carbon and LNMR), LNM and primary tumor is shown in Figure 4 (a-d). The first PCA scores plot (Figure 4a) demonstrates a separation of the malignant tissue types (primary tumor and LNM) from the non-malignant. This effect is demonstrated clearly by the second PCA scores plot (Figure 4c) where non-malignant LN tissue types are classified into one class. Leave-one-sample-out internal cross validation (Figure 5 a-e), with Mahalanobis distance classifier, demonstrated a classification accuracy of
96.7% for malignant tissue and 95.3% for non-malignant (Figure 5d). Receiver operating characteristic curve analysis demonstrated an AUC of 0.996 (Figure 5e). The 215 GPLs compared by means of statistical analysis in this data set are shown in Supplementary Table S2. With this proven association we were able to successfully implement our multivariate algorithm for the prediction of LNM based on an average primary tumor lipidomic profile of EA.

**Spatial prediction of LNM based on primary tumor lipidomic profiles**

The contingency table in Supplementary Table S3 demonstrates the full outcomes of our comparative analysis between the IHC reference test and the MSI Tissue Class Prediction Images (TCPI). In summary, the TCPIs had an overall sensitivity of 89.5%, specificity of 91.5%, positive predictive value of 73.9%, negative predictive value of 97.0% and an accuracy of 91.1% in diagnosing LNM when compared to the gold standard IHC. An example of matched IHC and TCPI images for four different types of lymph nodes is shown in Figure 6. The spatial distribution of the metastases is comparable between the two diagnostic modalities. Matched images for each of the 90 LNs are shown in Supplementary Figure S4. The single lymph node, which was classified as ITC on IHC had the presence of single/double red pixels on the TCPI. All LNs with the presence of LNMR were successfully classified as normal LNs on the TCPIs.

The main discrepancy in agreement between the reference and reference test was the presence of micro-metastases on the edge of 6 TCPIs but not on IHC. The IHC images do show positive AE1/AE3 staining in corresponding locations of these 6 LNS, however the histopathologists attributed this to non-specific edge effect of the sample. If we are also to exclude micro-metastases found exclusively on the edge of TCPIs, corresponding to the LN capsule, the results from the contingency table are markedly different (Supplementary Table
S4. In this case the TCPIs would have an overall sensitivity of 89.5%, specificity of 100%, positive predictive value of 100%, negative predictive value of 97.2% and an accuracy of 97.7% in diagnosing LNM when compared to the gold standard IHC. As metastases propagate in the sub-capsular sinus of LNs we would not expect to find them in the capsule, unless they are macro-metastases extending out from the parenchyma of the LN.

**Discussion**

In this study, the capability of DESI-MSI in identifying specific LN tissue types, with accurate spatial distribution and localization and without reliance on target specific reagents (e.g., antibodies) has been demonstrated. The quality of the DESI-MSI was suitable for the distinction of micro-metastases, which are defined by their maximum diameter of 0.2 to 2mm. Overall lipidomic profile analysis revealed a close association between the primary tumor and metastases in comparison to benign lymph node tissue types. This association was used for the objective identification of lymph node micro- and macro-metastases, based on the spatially resolved lipidomic profile of the tumor, with a sensitivity of 89.5%, specificity of 100%, positive predictive value of 100% and negative predictive value of 97.2% when compared to the gold-standard immunohistochemistry.

DESI-MSI may have a complementary role to histopathology that may overcome the limitations encountered in current practice. When the pathologist is unable to make a definitive diagnosis based on frozen section, DESI-MSI would have a supportive role allowing a definitive intraoperative diagnosis, in the same manner that special staining with immunohistochemistry supports paraffin embedded tissue diagnosis. Other potential applications for cancer diagnosis that would benefit from TCPI include identification of tumor type in metastases of unknown origin, accurate profiling of samples from fine needle aspiration or identification of tissue in histopathologically equivocal cases. In addition, it can be used for post-operative specimen analysis, which often requires extensive resources and
manpower to identify metastases in resected regional LNs. For instance, with esophageal cancer specimens, the number of resected LNs may approach 3 figures. An automated method may streamline the process and improve efficiency.

The following workflow illustrates a potential position of this technology in the cancer care pathway. A lipidomic profile of the primary tumor could be obtained pre-operatively through tissue biopsy or from a spectral database of that particular tumor type. This data, together with database entries obtained by the analysis of healthy lymph node tissues of other patients is then used to create a multivariate statistical model using the described algorithms for creating TCPI. The SLN can be sampled at the start of an operation and processed by a technician in an automated workflow as the surgeon continues to operate. Once the surgeon is informed about the results, he/she can make a decision regarding complete resection of the regional LNs. This method would objectify the process of LNM identification and reduce the burden on histopathologists as they do not need to be present at the time of the examination. Furthermore, analysis is not restricted to a single tumor type, which is the case for OSNA.

Previous studies report the use of other ionization techniques for the purpose of mass spectrometry imaging of LNM. The use of MALDI-MS has been demonstrated in studies of patients with melanoma, and breast cancer (36, 37). The MSI generated by MALDI primarily distinguishes areas of varying protein intensities, differentiating LNM from normal lymphocyte regions. Other groups of compounds such as trace elements, detected by laser ablation-inductively coupled plasma-mass spectrometry, have also been used to differentiate areas of metastases within LNs (38). DESI-MSI depends on similar principles of molecular differentiation, relying primarily on the detection of varying lipid signals in the 600-1000 m/z range. In comparison to other MSI techniques, DESI offers certain advantages with respect to clinical applications. It does not require matrix deposition (unlike MALDI) and since it works with full sensitivity under atmospheric conditions, it has the potential to be used in an
operating theatre with a portable mass spectrometer. Very similar lipidomic spectral features can be determined by technologies such as rapid evaporation ionization mass spectrometry (REIMS) (39,40), for in-vivo tissue identification of different types of lymph node/organ metastases. However, ex-vivo analysis with DESI-MSI may be more suitable for the identification of LNM, as mass spectra can be obtained from discrete areas at a microscopic rather than a macroscopic level. Furthermore, results can be corroborated by histopathological assessment post-operatively, as the tissue sections remain structurally unaltered by DESI-MSI.

Several studies comparing LNM with the primary tumor, have utilized ‘omic disciplines aside from metabolomic/lipidomic profiling. Perou et al. (41) demonstrated that the genomic profile of two lymph node metastases were similar to that of multiple primary breast tumor samples taken from the same patient. A further study of 26 patients with breast cancer used a hierarchical clustering method to group samples of primary tumor and paired metastases; it found that 92.3% of the cases clustered next to each other, indicating that their overall gene expression profiles were similar (42). The rate of proliferation of primary tumor and associated LNM has also been studied in 30 patients with breast cancer, by the identification of antibody labelled S-phase cells. Primary tumor and regional LNM labelling indices correlated strongly, not being influenced by age, level of hormone receptors, tumor size or number of positive nodes (43). In this study, the lipidomic profiles of the primary tumor were found to be similar to LNM. At the individual lipid level, two polyunsaturated PGs were found to be up-regulated in LNM. The functional outcome of these minor differences is undetermined but is likely to confer a survival advantage to cancer cells propagated in an environment foreign to the primary tumor.

Objective classification of LNM based on the average primary tumor lipidomic profile resulted in an accuracy of 97.7 %, which should encourage further investigation of this
method as a diagnostic test. The sensitivity and specificity of the technique are comparable to the current gold standard of immunohistochemistry without suffering from its subjectivity. In addition, the reproducibility of DESI-MSI for the analysis of human cancer tissue has been reported with a coefficient of variance of 18±8% (44).

Immunohistochemistry has historical precedence and has stood the test of time but suffers from the potential consequences of human error. The question that will remain unanswered is the true accuracy of the lipidomic profiling results of DESI-MSI, which may be compounded by the subjective nature of the reference test. Nevertheless, further independent prospective validation blind studies; pathway mapping and cost-effective analysis are needed to pave the way for the integration DESI-MSI in clinical practice.

Despite its translational potential, there are currently several limitations for DESI-MSI. In this study, the spatial resolution was 75µm. This is sufficient for identification of micro-metastases, which are defined as aggregations of cancer cells between 0.2 and 2mm in their widest dimension. However, identification of isolated tumor cells, (approximately 20µm in diameter) is limited at this level of resolution. The clinical advantage of identifying ITCs continues to be debated (45) and is thus not commonplace in the staging of most epithelial tumors. Nonetheless, an imaging resolution of 40µm has previously been described in the literature (46), and in our own institution we have employed DESI-MSI at a resolution of 20µm (Supplementary Figure S5). With the instrument parameters used in this research, the time required for imaging of very large lymph nodes was in the order of a few hours. Further development of this technology at our institution has improved the scan speed by a 30 fold increase without depreciation of image quality or information. This means that a 10 x 20 mm lymph node cross-section could be imaged with a resolution of 75µm in under 10 minutes (Supplementary Figure S6). The recent advances in spatial resolution and scan speed coupled
with the diagnostic potential of lipid biomarkers makes the translational application of DESI-MSI a distinct possibility.

**Conclusion**

DESI-MSI can be utilized for the objective identification of LNM based on its primary tumor lipidomic profile. The accuracy of 97.7% qualifies this technique as a diagnostic test for the identification of LNM independent of histopathological expertise. The use of lipidomic profiling in this context not only confers the advantage of analytical stability, but delivers a novel panorama of biomarkers that can be interrogated for diagnostic purposes.
References


### Table 1: Demographics and case characteristics of lymph node sample set.

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Demographics</th>
<th>Lymph nodes harvested for MSI</th>
<th>Histopathology of full resection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age</td>
<td>Sex</td>
<td>CTx</td>
</tr>
<tr>
<td>------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>TO22</td>
<td>68</td>
<td>M</td>
<td>Yes</td>
</tr>
<tr>
<td>TO26</td>
<td>56</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>TO28</td>
<td>78</td>
<td>M</td>
<td>No</td>
</tr>
<tr>
<td>TO29</td>
<td>59</td>
<td>M</td>
<td>Yes</td>
</tr>
<tr>
<td>TO30</td>
<td>80</td>
<td>M</td>
<td>Yes</td>
</tr>
<tr>
<td>TO31</td>
<td>62</td>
<td>M</td>
<td>Yes</td>
</tr>
<tr>
<td>TO32</td>
<td>57</td>
<td>M</td>
<td>Yes</td>
</tr>
<tr>
<td>TO39</td>
<td>38</td>
<td>M</td>
<td>Yes</td>
</tr>
<tr>
<td>TO40</td>
<td>50</td>
<td>M</td>
<td>Yes</td>
</tr>
<tr>
<td>TO41</td>
<td>66</td>
<td>M</td>
<td>No</td>
</tr>
<tr>
<td>TO42</td>
<td>55</td>
<td>F</td>
<td>Yes</td>
</tr>
</tbody>
</table>

CTx, Chemotherapy; LNs, Lymph nodes; Micro-Met, Micro-metastases; Macro-Met, Micro-metastases; LNMR, Lymph node metastases with response to Chemotherapy; T Stage, Classified as 1-4b as per 7th edition of TNM staging; N- Stage, classified as 1-3 as per 7th edition of TNM; LV, Lymphovascular Invasion. *LNMR was present in the same LN which contained the macro-metastases.
Figure legends

Figure 1: DESI-MSI of lymph node with macro-metastases and metastases response to chemotherapy (LNMR). a) AE1/AE3 immunohistochemistry image of the LN section post DESI-MSI; b) RMMC (supervised multivariate analysis) pixel classification image: metastases (red), LNMR (orange), lymph node connective tissue (green), lymph node parenchyma (blue); c) RMMC scores plot of tissue specific mass spectra in the 600-1000 m/z range; d) Principal component scores 1 image (PC1); e) Principal component scores 2 image (PC2); f) Representative mass spectra of tissue types; g) Heat map showing log₂ fold changes relative to the mean profile across all samples.

Figure 2: DESI-MSI of lymph node with micro-metastases. a) AE1/AE3 immunohistochemistry image of the LN section post DESI-MSI; b) RMMC (supervised multivariate analysis) pixel classification image: metastases (red), lymph node connective tissue (green), lymph node parenchyma (blue); c) RMMC scores plot of mass spectra in the 600-1000 m/z range; d) zoomed in image of micro-metastases, e) Principal component scores 3 image (PC3); f) Representative mass spectra of tissue types; g) Heat map showing log₂ fold changes relative to the mean profile across all samples.

Figure 3: Multivariate analyses comparing lipidomic profiles (m/z 600-1000) of LNM (red) versus primary tumor (green) and LNM versus LNMR (blue). a/c) Principal component analysis scores plot (each point is the average of multiple mass spectra representative of that tissue type within a single sample from one patient); b/d) representative mass spectra of LNM, primary tumor and LNMR.

Figure 4: Unsupervised multivariate analyses comparing lipidomic profiles (m/z 600-1000) of specific non-malignant lymph node tissue types versus metastases and primary tumor. a) Principal components analysis scores plot (each point is the average of multiple mass spectra
representative of that tissue type within a single sample from one patient); b) Average mass spectral profiles of specific tissue types in full dataset; c) Principal components analysis scores plot comparing lipidomic profiles of non-malignant lymph node tissue types (grouped as one class) versus metastases and primary tumor; d) Average mass spectral profiles of non-malignant LN tissue, metastases and primary tumor.

**Figure 5:** Internal cross validation – lipidomic profile of malignant tissue (tumor and LNM) versus non-malignant LN tissue types (LNP, LNC, fat, carbon). a) Principal components analysis scores plot (each point is the average of multiple mass spectra representative of that tissue type within a single sample from one patient); b) RMMC scores plot; d) Leave-one-sample-out cross-validated RMMC scores plot as per confusion matrix; d) Confusion matrix of leave-one-sample-out internal cross validation with Mahalanobis distance classifier; e) Receiver operating characteristic curve.

**Figure 6:** Examples of matched immunohistochemistry (IHC) and Tissue Classification Pixel Images (TCPI)
Component 1 (45.6%)

Metastases Response to CTx

Lymph node connective tissue

Lymph node parenchyma

Metastases

Log fold change relative to the mean profile across all samples

Figure 1
Figure 2

a. [Image of lymph node connective tissue]
b. [Scatter plot showing component 1 and 2]
c. [Box plot showing log fold change]
d. [Image of lymph node parenchyma]
e. [Image of micro-metastases]
f. [Bar chart showing intensity (a.u.)]
g. [Heatmap showing data points across PC1 and PC3]
Figure 3

(a) PCA plot with Component 1 (20.0%) and Component 2 (14.3%).
(b) Mass spectrum with m/z values.

(c) PCA plot with Component 1 (20.3%), Component 2 (16.9%), and Component 3 (16.4%).
(d) Mass spectrum with m/z values.
Figure 4

PCA

Component 1 (20.5%)
Component 2 (12.3%)
Component 3 (10.0%)

Comp. 1 (21.8%)

PCA

Component 2 (13.7%)
Component 1 (20.5%)

Carbon
Fat
LNC
LNP
LNM
LNMR
Primary Tumour

Intensity (a.u.)
m/z

Non-malignant LN tissue
LNM
Primary Tumour

Intensity (a.u.)
m/z
Component 1 (14.1%)  
Component 2 (0.0%)

- **Malignant**
- **NonMalignant**
- Incorrectly classified into Malignant group
- Incorrectly classified into NonMalignant group

**True Class**  
96.7%, 29  
4.7%, 3

**Predicted Class**  
3.3%, 1  
95.3%, 61

**ROC curve, AUC = 0.996**

**Figure 5**
Macro-metastases

Normal LN with evidence of carbon deposits

Micro-metastases

Metastases with complete response to CTX