Differential Mobility Spectrometry Imaging for Pathological Applications

Anton Kontunen[a]*, Jalmari Tuominen[b]*, Markus Karjalainen[a], Osmo Anttalainen[b], Teemu Tolonen[c], Pekka Kumpulainen[b], Maiju Lepomäki[b], Antti Vehkaoja[a], Niku Oksala[b,d], Antti Roine[b]

[a]Faculty of Medicine and Health Technology, Tampere University, Korkeakoulunkatu 7, 33720 Tampere, Finland

[b]Department of Surgery, Faculty of Medicine and Health Technology, Tampere University, Korkeakoulunkatu 7, 33720 Tampere, Finland

[c]Department of Pathology, Fimlab Laboratories and Tampere University Hospital, Arvo Ylpön katu 4, 33520 Tampere, Finland

[d]Division of Vascular Surgery, Tampere University Hospital, Elämänaukio, Kuntokatu 2, 33520 Tampere, Finland

*Authors contributed equally

Correspondence

Anton Kontunen, Faculty of Medicine and Health Technology, Tampere University, Korkeakoulunkatu 7, 33720 Tampere, Finland. Email: anton.kontunen@tuni.fi

Abbreviations[1]

[1] MSI = mass spectrometry imaging; DMS = differential mobility spectrometry; MALDI = matrix-assisted laser desorption ionization; SIMS = secondary ion mass spectrometry; DESI = desorption electrospray ionization; REIMS = rapid evaporative ionization mass spectrometry; ATAS = automatic tissue analysis system; CS = chemical stress; LDA = linear discriminant analysis; 10-f-CV = 10-fold cross-validated; FSFS = forward sequential feature selection

Abstract

Pathologic examination of clinical tissue samples is time consuming and often does not involve the comprehensive analysis of the whole specimen. Automated tissue analysis systems have potential to make the workflow of a pathologist more efficient and to support in clinical decision-making. So far, these systems have been based on application of mass spectrometry imaging (MSI). MSI provides high fidelity and the results in tissue identification are promising. However, the high cost and need for maintenance limit the adoption of MSI in the clinical setting. Thus, there is a need for new innovations in the field of pathological tissue imaging. In this study, we show that differential ion mobility spectrometry (DMS) is a viable option in tissue imaging. We demonstrate that a DMS-driven solution performs with up to 92% accuracy in differentiating between two grossly distinct animal tissues. In addition, our model is able to classify the correct tissue with 81% accuracy in an eight-class setting. The DMS-based system is a significant innovation in a field dominated by mass-spectrometry-based solutions. By developing the presented platform further, DMS technology could be a cost-effective and helpful tool for automated pathological analysis.

Keywords: differential ion mobility spectrometry, diathermy, imaging, mass spectrometry, tissue mapping

Introduction

The aim of oncologic surgery is to remove the tumour with adequate healthy tissue margin (Orosco et al., 2018). The assessment of margins and precise histological diagnosis is also the main goal of a pathologist. Since the processing of the tissue specimen as well as the microscopic assessment are largely manual processes, the capacity of a single pathologist is limited, and it is impractical to analyse the removed tumour tissue completely. Instead, the pathologic-anatomical diagnosis is set by investigating only selected sections of the tissue, risking false negatives, especially in intraoperative frozen section analysis and imprint cytology (Esbona et al., 2012; Tucker, 2012).

In order to increase the accuracy and to streamline the work of the pathologist, image recognition solutions have been developed for screening the histologic slides (Komura and Ishikawa, 2018; Nature, 2017). Although these solutions clearly hold promise, they rely on conventional fixation and staining and thus potentially automate only a part of the workflow. A method capable of analysing tissue with minimal preparation would allow efficient and complete assessment of histological specimens without unacceptable overhead for the pathology laboratory. Such a method would be especially useful as an orthogonal supporting method for histopathology rather than its replacement. The results gained from a more spatially comprehensive, but potentially less accurate analysis could be used to pinpoint areas for a more specific microscopic examination. This could be achieved by creating a so-called tissue map that highlights the points of interest or types of tissues before the gold standard analysis.

Tissue mapping with molecular methods is an approach that has seen extensive use in basic science and pharmacologic research. These solutions mostly rely on Mass spectrometry imaging (MSI), more specifically on matrix-assisted laser desorption ionization (MALDI), secondary ion mass spectrometry (SIMS), desorption electrospray ionization (DESI), and rapid evaporative ionization mass spectrometry (REIMS) (Golf et al., 2015; Veselkov et al., 2014). These methods have been shown to be accurate in tissue identification and a widespread adoption in clinical setting has been speculated (Veselkov et al., 2014). Except for MALDI, these systems do not require extensive pre-processing of the samples and provide molecular data that may be easier for machine learning algorithms to utilize than complex histologic views (Golf et al., 2015; Veselkov et al., 2014). However, the adoption rate of MSI is limited by the high cost of the systems and their upkeep. More

cost-effective analysis methods are required before MSI could be considered for widespread clinical use. (Golf et al., 2015; Veselkov et al., 2014)

Differential ion mobility spectrometry (DMS) is an analytical technique related to mass spectrometry. The key difference is that DMS does not require a high vacuum to function, although many commercial sensors that are primarily intended to act as a pre-filtering stage for mass spectrometry or mobile use, operate slightly below atmospheric pressure to improve signal intensity and to reduce the energy consumption of the system (Nazarov et al., 2006). However, some stand-alone DMS sensors, such as the one used in this study, can even operate in atmospheric pressure, resulting in good reliability, ease of maintenance, more robust and simpler sampling, and lower costs, while maintaining sufficient performance. Ion mobility spectrometry has been shown to have various applications in the medical field (Chouinard et al., 2016; Covington et al., 2015; Kabir and Donald, 2017), but its use has not yet spread to pathology. We have previously shown that DMS coupled with a diathermy blade is a viable technology in tissue identification (Kontunen et al., 2018; Sutinen et al., 2019). In our studies, we were able to identify various porcine tissues with over 95% classification accuracy and human breast malignancies and control tissues with over 85% accuracy. The automatic tissue analysis system (ATAS) was developed to automate sample analysis and data collection. However, the previous systems were not suitable for high-throughput analysis due to slow sampling rate. We modified the system to enable higher throughput and resolution. Additionally, we created a method for the generation of topographical data that enables mapping of specimens in a manner analogous to large-format histology. As a relatively inexpensive technology, DMS has the potential to bring molecular tissue imaging to wider use.

In this proof-of-concept study, we present a cost-efficient system for automated histopathological tissue mapping using DMS technology and machine learning. We also demonstrate, how the visual output data of the analysis system could be superimposed on a tissue image in a future pathological application.

Materials and Methods

Fresh tissue samples of a Finnish landrace pig *(Sus scrofa domesticus)* were obtained from a slaughterhouse (Paijan tilateurastamo, Urjala, Finland) and immediately stored in a freezer at -18 °C. The tissues included in this study were lungs, kidneys, liver, brain, and flank that had clear areas skeletal muscle and adipose tissue.

The brains that were used in the measurements were gathered from 5 pigs, but all other tissues came from a single animal. The tissues were slaughterhouse offal and commercial meat products. Therefore, no approval from the ethical committee was needed for the study. Prior to analysis, the tissue samples were cryosectioned into 3 mm thick sample slices with a meat slicing machine (Prego P119, Inbound, Finland) and placed on a protective agar plate that produces no significant signal, even if the diathermy blade evaporates part of it.

ATAS

The agar plate was positioned on the sampling platform of the ATAS system. The platform consisted of a modified 3D printer, in which the printer bed functioned as the dispersive electrode for the diathermy knife that was controlled by the graphical user interface of the printer. The sampling platform along with the subsequent parts of the ATAS system have been used and described in detail in our previous studies (Kontunen et al., 2018; Sutinen et al., 2019). In short, a 3D-printer-mounted diathermy blade is used to cut a tissue sample resulting in surgical smoke that is filtered and diluted in the sample pre-processing unit, after which it is measured with the DMS. The DMS device used in this study was the ENVI-AMC DMS sensor (Environics Oy, Finland), which is capable of detecting low parts-per-billion concentrations of gaseous substances. The device is 178 mm high, 440 mm wide, and 517 mm deep, which equals to roughly the same size as a standard surgical smoke evacuator. The sensor separates the sample ions by an asymmetric square waveform that has a duty cycle of 5% and frequency of 250 kHz. The strength of the separation field can be varied between 10 to 80 Townsend (Td) with a resolution of less than 0.01 Td. A more in-depth description of the technical details of the DMS sensor can be found in a dedicated publication (Anttalainen et al., 2018).

A schematic representation of the ATAS used in this study is shown in Figure 1. The main difference compared to the previous system (Sutinen et al., 2019) is that the diathermy blade was modified to a needle shape to allow for increased spatial resolution. There was also no sample chamber inside the sample pre-processing unit that would accommodate a longer measurement duration for the DMS device. The DMS was set to sweep the compensation voltage (x-axis in the output dispersion plot) from -0.8 V to 5.0 V with 25 increments and the waveform amplitude voltage (y-axis in the output dispersion plot) from 340 V to 740 V with 8 increments. This resulted in a more sensible timeframe for the imaging of the full tissue slice, but also in a relatively low DMS measurement resolution (25x8 = 200 pixels) for each sample point in the tissue matrix.

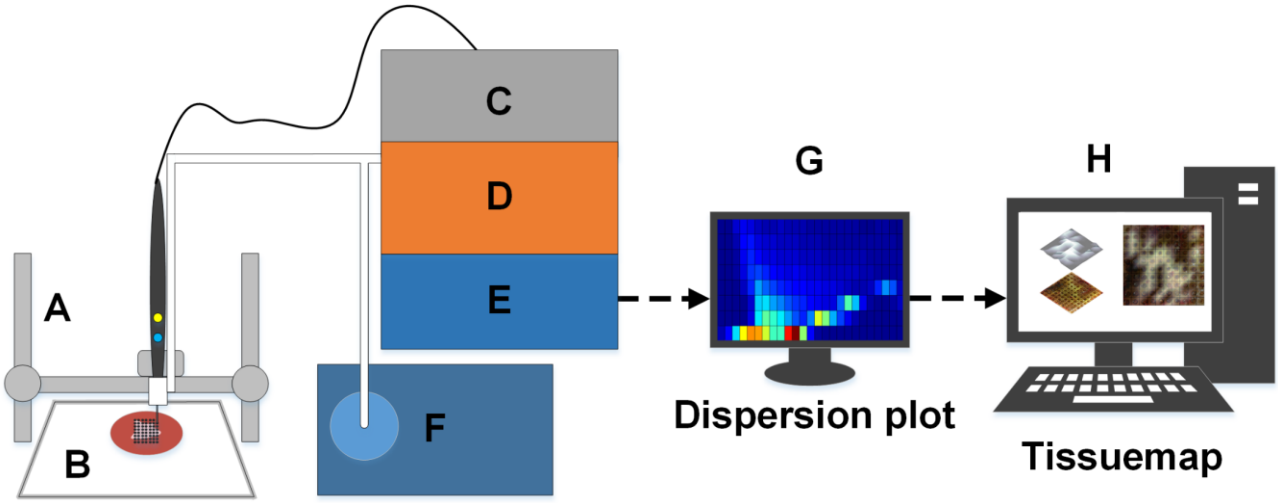

Figure 1. The automatic tissue analysis system (ATAS) system for tissue imaging in a pathological application. A) 3D printer-based sampling unit. B) Tissue slice on the sampling platform. C) Standard diathermy unit. D) Sample pre-processing unit. E) ENVI-AMC differential mobility spectrometer (DMS). F) Surgical vacuum. G) Monitor showing the dispersion plot of a sample point. H) Computer for data-analysis and visualization of the DMS data.

The dimensions of a tissue matrix were set with the graphical user interface of the printer. Generally, the ATAS was set to analyse the tissue slice with a 1–2 mm spatial resolution between each sample point in x and y directions. The incision depth of the needle blade was set to 3 mm, so that the depth of the cut would match or only slightly exceed the full depth of the tissue slice. To prevent carry-over signal from the cut of the preceding sample point, a waiting duration of 10–30 seconds after each cut, was also set. The waiting duration was dependent on the analysed tissue type, with liver requiring the longest time and muscle the shortest time. The waiting times were empirically set, so that the accumulation of the surgical smoke signal was not visible in the spectrum regardless of the number of consecutive cuts. With these measurement settings, a matrix consisting of 100 sample points was analysed by the ATAS in approximately 25–75 min.

Measurement Protocol

Prior to analysis, the tissue samples were photographed from a distance of approximately 10 cm using a macro lens (HD MACRO, BLACKEYE, Eye Caramba Oy, Finland) attached to a mobile phone (iPhone 6, Apple Inc, USA). After performing the measurement set to produce a tissue matrix as previously described, the sample was photographed again (Figure 2). These photographs were used to annotate each pixel of the tissue matrix. Due to the size of the needle blade, some cuts evaporated the tissue from areas where the tissue type was difficult to determine macroscopically. Thus, only a portion of the points were confidently annotatable,

and these points were independently marked as such. Clear macroscopic histological heterogeneity was observed in kidneys, brains and flank. The tissues therein were identified as renal cortex and renal pelvis in the kidneys; grey matter and white matter in the brains; and skeletal muscle and intermuscular fat in the flank. All cuts to liver and lung tissues were annotated as such due to their macroscopic homogeneity. All tissues were measured during a period of 1.5 months, with emphasis on minimizing any sources of bias produced by the measurement protocol. This means that the tissue types were measured on several different days and in varying consecutive orders.

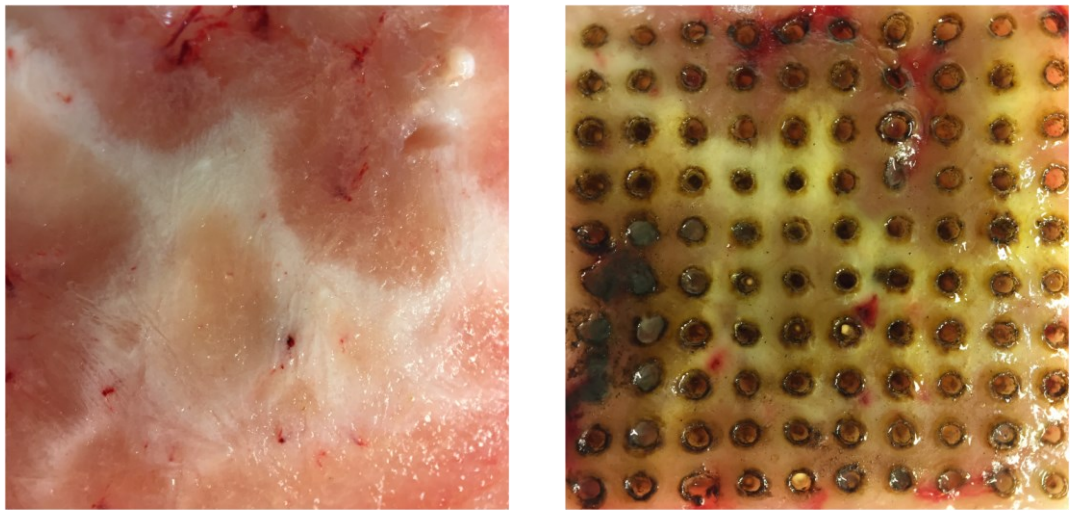

Figure 2. Examples of images used to annotate tissues. A picture taken from a porcine brain sample before automatic tissue analysis system (ATAS) tissue mapping (left) and after the full matrix of 100 sample points with 2 mm spatial resolution for additional clarity (right).

Data analysis and statistical tools

Statistical analyses and classifications were primarily done in MATLAB (version R2017b, The MathWorks Inc., Natick, USA). Additional models were also tested with a number of open source packages written for Python (version 3.6.6). The Python classification algorithms were implemented using scikit-learn (version 0.20.0) (Pedregosa et al., 2011). Visualizations were created using MatPlotLib (version 2.2.2) and Seaborn (version 0.9.0). Exploratory analyses were performed in Jupyter Notebooks.

Data

The complete dataset consists of 4742 DMS spectra in vector form, which has 200 dimensions, where each element is a 16-bit value that signifies the signal output of the ion detector with the corresponding DMS voltages. Each DMS spectrum represents one point in a tissue analysis matrix. Based on the DMS spectra, we derived a novel parameter, chemical stress (CS), to monitor the signal intensities of the DMS measurements. The CS value was defined as the quotient of the amount of heavy volatile organic compounds (i.e. signal in the bottom left of a DMS dispersion plot) to the amount of reactant ions (i.e. signal in the bottom centre of a DMS dispersion plot). A CS value of 3.0 was used as a cut-off value differentiating a saturated sample from a non-saturated one. In addition to the saturated samples, a software defect in the ENVI-AMC spectrometer resulted in a number of corrupted spectra that could not be used in the result analysis. Furthermore, in order to ensure the quality of the labelling, all uncertain labels were excluded, resulting in the final dataset of 3418 samples, which was used to train and validate the classification models.

Data pre-processing

Before configuring and training the classification models, the raw data vectors from each sample were adjusted to have only positive real numbers by offset correction. In addition, a common logarithm was calculated of every dimension. The logarithmic representation of the data was done to emphasize the information otherwise lost due to the prominence of high-intensity peaks in the ion spectrum. In other words, the logarithmic pre-processing evened out the disproportion of signal intensity changes between the DMS spectra and made the smaller changes in the key parts of the spectra more pronounced. After the pre-processing steps, the data were analysed in three phases.

Analysis

Phase I: *Internal classification and clustering*

The most important future application of the ATAS is automated tumour margin detection as previously described. Hence, we investigated to which extent our system is able to distinguish different tissues within the tissue slice samples. Macroscopically distinguishable tissue types were easy to identify in the flank, kidneys and brain, and thus these tissues were investigated in phase I.

The classification algorithms in this phase were based on linear discriminant analysis (LDA) models that have also been successfully utilized in previous studies (Kontunen et al., 2018; Sutinen et al., 2019). Models for the internal classification of flank, kidneys, and brain were created by utilizing the corresponding subgroups of the full data set that had only the sample points of each target tissue. For the internal classification of flank, the size of the data was 1240 samples (721 samples of skeletal muscle and 519 samples of intermuscular fat); for kidneys 822 samples (631 samples of renal cortex and 191 samples of renal pelvis); and for brains 470 samples (227 samples of grey matter and 243 samples of white matter). All models were 10-fold cross-validated (10-f-CV) to alleviate overfitting.

In addition to the LDA classification, K-means clustering was also used in the data analysis for individual tissue slices, to see if an unsupervised method could be used in DMS-based tissue mapping. K-means is a clustering method, where the algorithm tries to iteratively find centres of data groups with the shortest (Euclidian) distances to the data points surrounding it, when K, the number of the groups, is known (Jin and Han, 2010). The K-means clustering could be a well-suited method to produce tissue maps without prior training or annotation of any of the sample points in the tissue matrix.

Phase II: 8-class classification

Even though tissue imaging by unsupervised K-means clustering might be possible, supervised classification by a pre-trained model most often yields better results. In order to test the limits of our classification system, we aimed to create a model that is, after training, able to identify any of the used tissues without prior knowledge of the sample. The model was trained by using the full dataset of 3418 samples from the 8 tissue types and the classification method used was LDA with 10-f-CV.

Phase III: Identifying key features for classification

The key features of the ion spectrum for the classification of the tissues were searched by forward sequential feature selection (FSFS). The function of FSFS is to determine the most relevant pixels of the dispersion plot in the classification of certain groups. The selection of the relevant features starts with a blank vector, for which the algorithm first adds the feature that yields the best classification result. After this, the FSFS algorithm continues to search for features that further improve the result, until a stopping criterion is met

(Wang et al., 2017). In practice, this means that the FSFS algorithm adds features to be used in the classification, until the addition of new features no longer improves the classification result significantly. The FSFS process was executed by classifying each target tissue in a binary setting against all other tissues, in order to find the significant features that explain their differentiation with LDA classification. Since there is some variation in the DMS outputs of even the same tissues, it is not certain that the results of a single FSFS are generalizable to represent the features of the tissue conclusively. In other words, the results are affected by the structure of the training data, which changes if cross-validation is used. Thus, the FSFS process was repeated 1000 times for each tissue, to trace the features that are not heavily dependent on the cross-validation partitions (i.e. chosen in the majority of FSFS cycles).

Results

Internal classification and clustering

The internal differentiation of tissues in flank, kidneys, and brains was tested with the unsupervised 2-means clustering and the trained LDA models. The binary K-means clustering results for a porcine flank tissue matrix can be seen in Figure 3, along with example dispersion plots for skeletal muscle and intermuscular fat. It is apparent from the figure that while the K-means clustering can differentiate the majority of the areas with intermuscular fat and skeletal muscle, the produced tissue map is not perfect. The mapping results were similar for all tissue slices, indicating that the pathological mapping with unsupervised clustering of DMS data is not yet a completely viable option.

With the supervised LDA models, the cross-validated accuracy score (mean and standard deviation of 100 CV repetitions) for intermuscular fat and skeletal muscle was 91.8% ± 0.3% (n=1240), 70.9% ± 1.5% (n=470) for white and grey matter, and 90.8% ± 0.5% (n=822) for renal cortex and renal pelvis.

The models trained for internal classification of porcine flank, kidney or brains, can be used to create a more accurate map of the sample points of a new tissue slice, than with unsupervised methods. The posterior probability values that the LDA classifier produces, can be used to represent the certainty of the classification and can be visualized as a color-coded heatmap, where each pixel represents one sample point in the tissue matrix. While this method has little value with the tissues that have clear macroscopic areas that can be

identified with the bare eye, the method could be utilized to guide a more focused sampling in pathological screening of clinically relevant tissues, where the visually apparent differences are not as profound. Furthermore, in a potential future application, the classification heatmap could be overlaid on top of an RGB-image of the tissue slice, thus making the interpretation of the areas even more intuitive for a pathologist performing histological analysis. This idea is demonstrated in Figure 4.

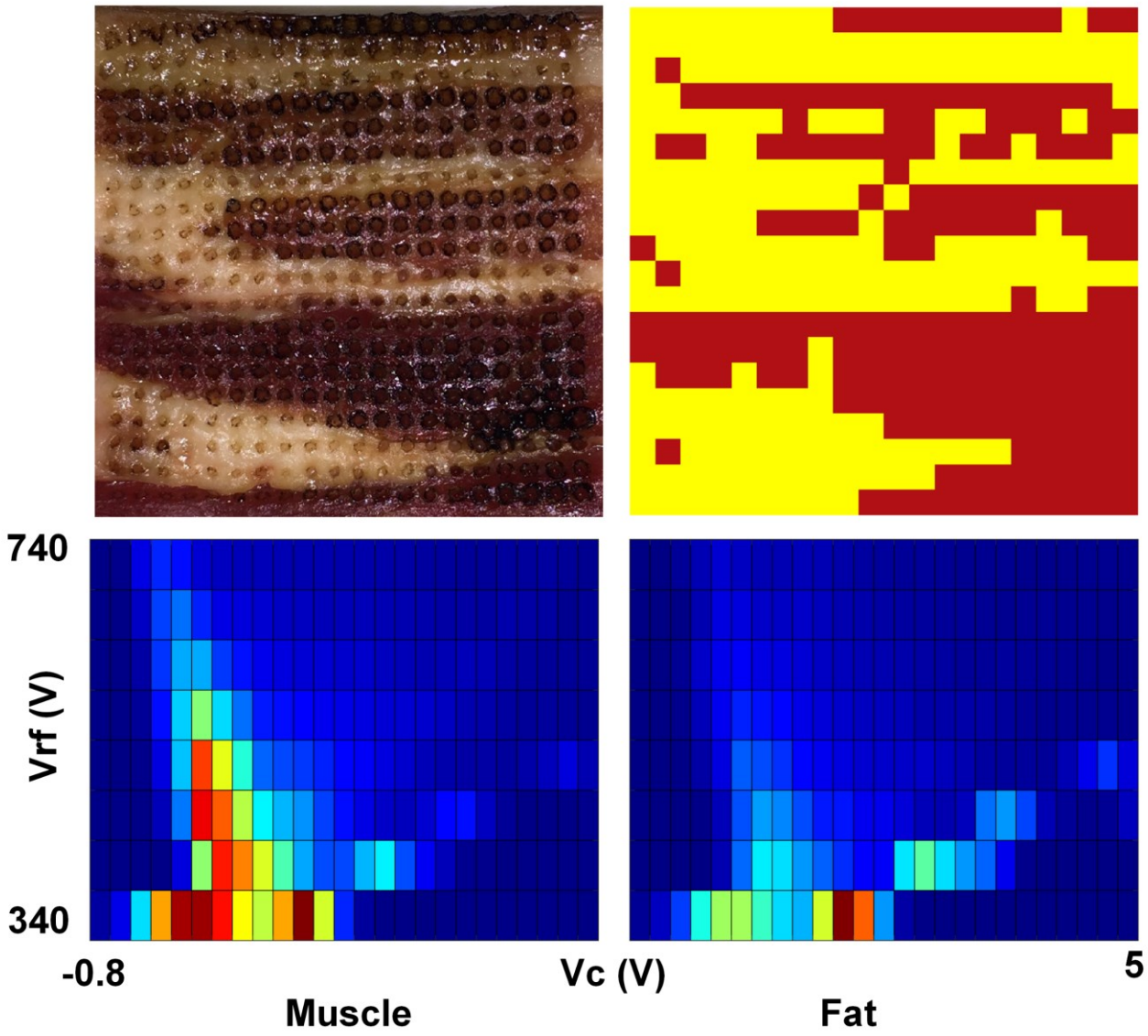

Figure 3. K-means clustering for skeletal muscle (red) and intermuscular fat (yellow) and an example dispersion plot from both tissues.

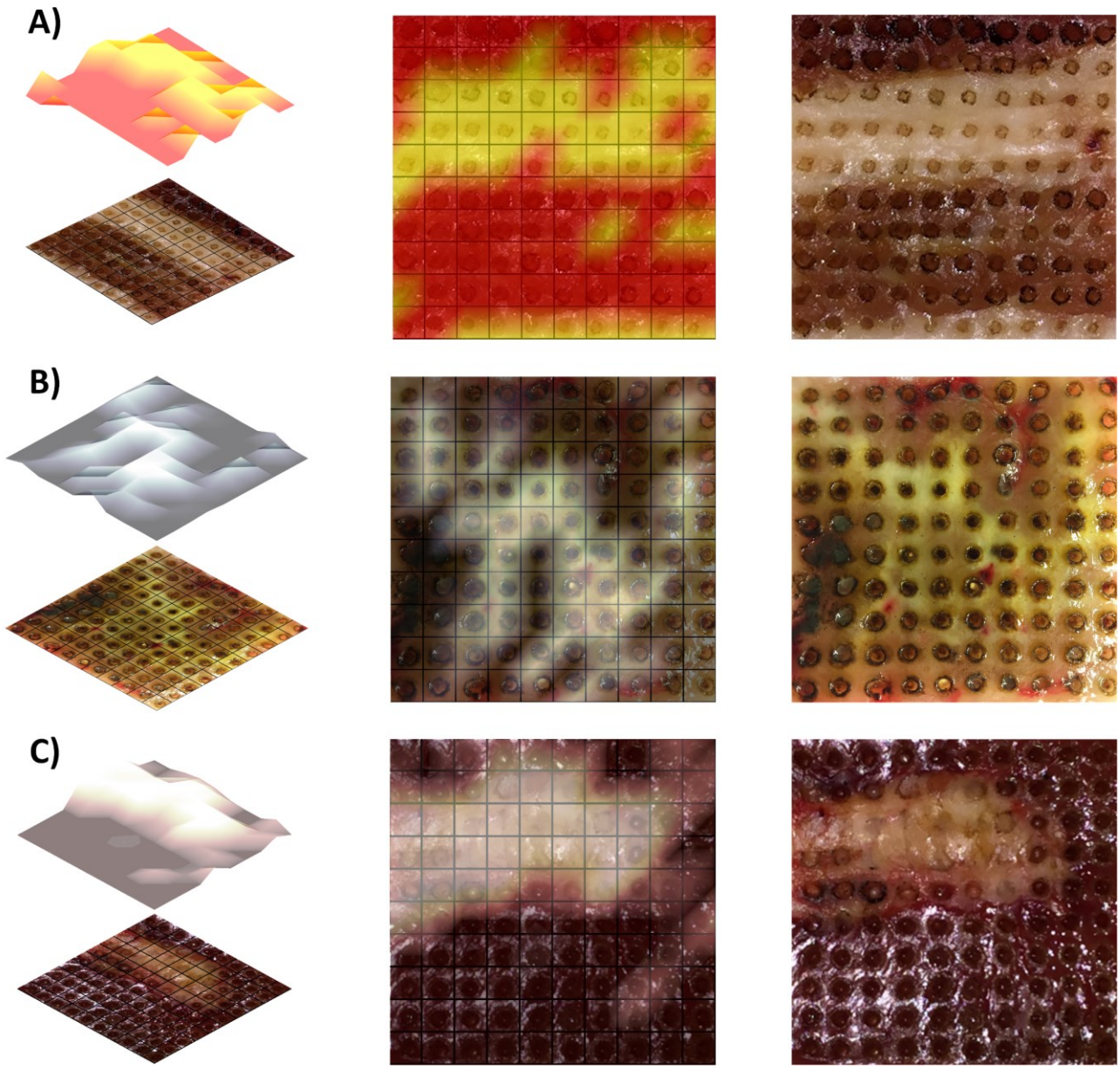

Figure 4. Linear discriminant analysis (LDA) classification heatmap projected on top of an imaged tissue slice with 100 sample points (10x10 matrix with 2 mm spatial resolution for clarity). A) Porcine flank sample. The heatmap illustrates the classification certainty of each sample point. A completely red pixel indicates that the model classifies the sample point as muscle with 100% certainty, whereas a completely yellow pixel indicates the equivalent for fat. B) Porcine brain sample, where the heatmap illustrates areas of white matter (white), and grey matter (black). C) Porcine kidney sample, where the heatmap illustrates areas of renal pelvis (white), and renal cortex (dark brown).

8-class classification

The eight tissues were identified with a mean accuracy score of 81.4% ± 0.2% (n=3418). Skeletal muscle was correctly classified in approximately 90% of the cases, whereas only 72% of renal pelvis was correctly identified. The classification accuracy of the rest of the tissues fall between these ranges. Detailed results are presented as a confusion matrix in Table 1.

Table 1. A confusion matrix for a 10-fold cross-validated linear discriminant analysis (LDA) classification model. SM = skeletal muscle, IF = intermuscular fat, RC = renal cortex, RP = renal pelvis, GM = grey matter, WM = white matter. Rows represent the true class and columns represent the predicted class.

	SM	IF	RC	RP	GM	WM	Lung	Liver
SM	**584**	**39**	**34**	**11**	0	1	**45**	7
IF	47	**438**	7	6	0	0	**17**	4
RC	9	2	**542**	**27**	4	1	**19**	**27**
RP	1	**10**	**46**	**128**	0	1	3	2
GM	0	0	3	0	**155**	**58**	2	9
WM	0	8	5	0	**53**	**168**	0	9
Lung	6	8	**59**	5	1	0	**416**	4
Liver	0	4	**27**	1	0	1	2	**352**

Identification of key features for classification

The results of the FSFS for each tissue type indicated that the mobility spectra of the tissue smoke contain certain key features that the classification algorithms can use to identify the type. The type-specific features for each tissue are presented in Figure 5. The number of features found by FSFS was dependent on the tissue being classified, but on average, 8 features were found until the stopping criterion was met.

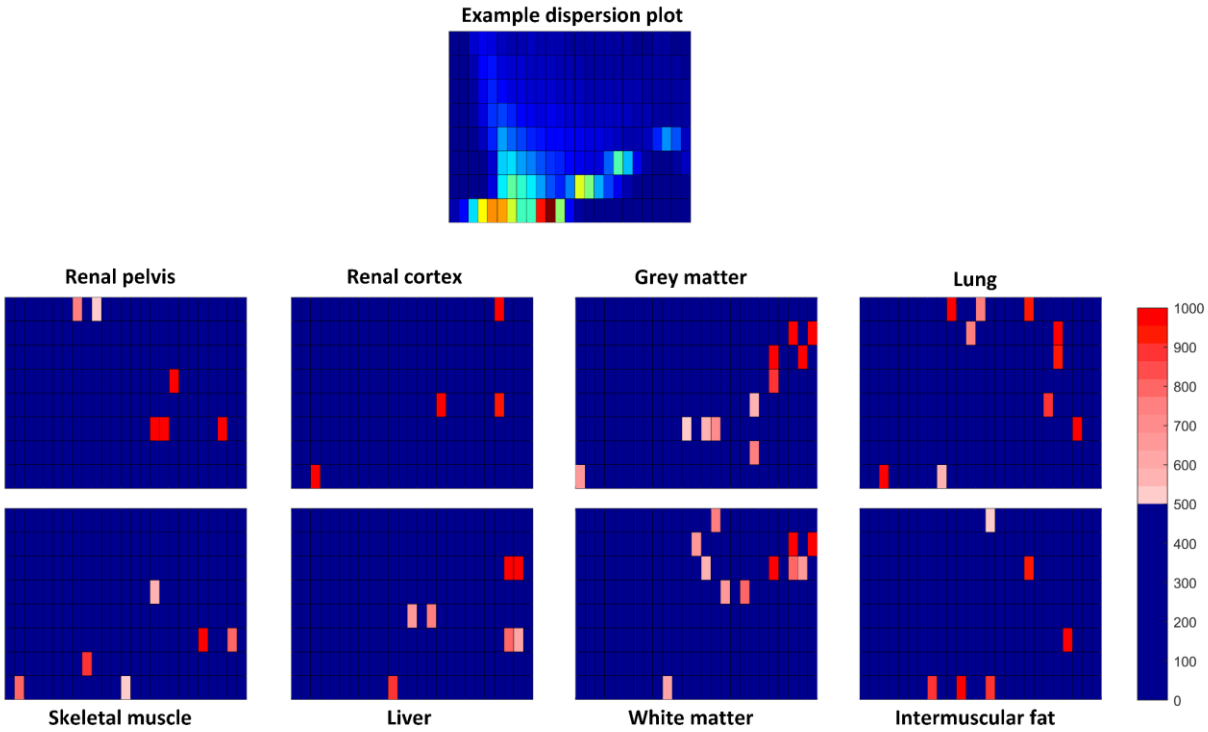

Figure 5. Representative dispersion plot and the results of the forward sequential feature selection (FSFS) for the identification of the used tissues. The red areas of the lower dispersion plots indicate the key features that the linear discriminant analysis (LDA) classification model uses to differentiate each tissue in a binary setting against all other tissues. The shade of red represents the number of times (out of a thousand) that the pixel was a result of the feature selection process. All pixels that were identified as key features in less than half of the FSFS cycles are blue.

Discussion

In this study we demonstrated a DMS-based automated tissue analysis system that is capable of topographic mapping of tissue specimens. With further development, this technology could potentially support pathologists in their analysis and decision-making. We achieved sufficient spatial resolution for tissue types which would also allow the detection of margins of a macroscopic tumour. It is noteworthy that these results were gained with a prototype system that can be significantly improved in terms of analysis time and resolution with future iterations of the device. In a future application, the tissue map resulting from the ATAS analysis could guide a pathologist to take specific areas under closer microscopic assessment. This would require two planar samples from the same area, due to the destructiveness of the analysis method, which is also the case with other current molecular mapping methods, such as MALDI, DESI, or REIMS.

The classification accuracy did not reach the results obtained in prior MSI-driven studies. However, this was expected, given the simplicity of our system. Even in an intraoperative setting with clinically relevant tissues,

REIMS identification studies have produced classification accuracies that are consistently over 90% (Balog et al., 2013; Phelps et al., 2018; St John et al., 2017). Furthermore, the results and figures presented in the article describing the pathological application, indicate classification accuracies of 100% with animal tissues (Golf et al., 2015). The difference between our results and the MSI studies is likely partially explained by technical factors. The specimens exhibited some variation and nonuniformity in the height and surface flatness that in turn led to variations in signal smoke concentration and signal intensities that were not yet compensated in the measurement protocol or system structure. Additionally, day-to-day variation was apparent in the measurements that could be explained by changes in environment factors such as ambient air temperature and humidity. Furthermore, mass spectrometers have higher resolving power that likely enables them to detect features that are not sufficiently separated by DMS.

In practice, the performance difference of classification accuracy of 81% as compared to 95-100% achieved by mass spectrometry systems means that in its current form, the system would require a confirmative test, such as a histologic slide. Since the classification algorithm can be weighted to prioritize high sensitivity or specificity by sacrificing the other, it would be worthwhile to tailor it to meet the demands of the application. In case of cancer specimens, it would be useful to weight the algorithm for high negative likelihood for cancer where it could be used to screen the specimens that are currently omitted from histological assessment due to resource constraints.

The intensity of the evaporated tissue matter proved to be one of the most significant factors in tissue classification. In addition, we observed a daily drift in the intensity of the measurements that we compensated by adjusting the dilution of the sample flow. The goal was to prevent the saturation of measurements, since the DMS responses became identical (i.e. saturated) in high signal intensities. These problems require further research in order to 1) uncover the cause of daily drift in signal intensity, 2) discover the optimal threshold of signal intensity and 3) create a dynamic dilution system resulting in optimal tissue differentiation preserving the intrinsic signal intensity of a given tissue.

Besides the intensity of the measurements, z-directional tissue overlap potentially affected the results. The blade was set to cut 3 mm in depth in order to produce enough surgical smoke for the detector. The suboptimal classification results in the brain are most likely explained by the overlapping of cortical tissue; it is likely that

during the 3 mm vertical movement, both grey and white matter were evaporated and detected by the sensor. This indicates that the method is best suited for planarly uniform tissue samples. Same extent of overlapping is not present in the kidneys or the flank but may explain some misclassifications within these subgroups.

We were able to identify features of the spectrum with FSFS that differentiate the tissues with relatively high accuracy despite the long-term drift and variation within tissue matrices. Certain features were always present, while others appeared only in a subset of spectrums indicating overfitting rather than a true discriminative feature. Some of the found features reside in the upper area of the DMS dispersion plot that has a lower signal-to-noise ratio. The noise in these areas may have a significant detrimental effect on classification performance and may explain part of the misclassifications.

The average number of features found with FSFS was 8, and the maximum number of features in any of the cycles was 39. Given that the full dispersion spectrum has 200 pixels, the distinguishing features account for less than 20% of the spectrum. We speculate that ATAS-based tissue imaging could be accelerated significantly by focusing the DMS sweep only on areas with high importance in tissue identification. However, this is not possible with the DMS device used in this study or any other commercially available DMS.

Limitations

Although the overall classification results were promising, a significant day-to-day bias was discovered. Even though we strived to eliminate all confounding factors, we were unable to completely control the ambient humidity or changes in the compressed air network of the laboratory. This highlights the need to introduce controls such as calibration to improve the repeatability of the system and to mitigate the external factors that affect the output of the DMS device.

In addition to the changes in the DMS output dispersion plots over time, it is possible that the Linear Discriminant model suffers from overfitting resulting in optimistic classification accuracy. It is possible that, the tissue-wise classification results are affected by daily variations in the measurement conditions to some extent. However, because the sample randomization and circulation between measurement days were taken into consideration, it is more likely that the bias produced by daily variations is rather negative than positive in this study. This is also evident from the fact that we failed to achieve the classification accuracy of our

previous study, where we identified 10 different porcine tissues with 95% classification accuracy (Kontunen et al., 2018). In the previous porcine tissue article, we speculated that the results of the second phase of the study might be overly optimistic, since the measurement protocol in that particular phase was not designed optimally and enabled day-to-day bias to some extent. The results gained in this study with better study design and varying sample material between all measurement days, indicate that our previous results were indeed partly biased.

Even though the study design in this research was properly executed, the system was not fully optimized to prevent carry-over, which led to relatively long delays that could be shortened by simple modifications to the system design. By minimizing the tubing lengths in all parts of the system and by optimizing the smoke sampling process, the creation of an even larger tissue matrix in significantly less time would be feasible.

The maximal resolution of the tissue matrices was limited by physical dimensions of the needle-like cutting blade. Although the imaging resolution of 1 mm is sufficient in postoperative margin detection, it is naturally unable to detect microscopic tumours of smaller diameter. This limits the diagnostic capability of the system and is a target for improvement. In order to improve the spatial resolution, the tissue evaporation could be done with a high-power laser, which would enable µm-scale distances between the sample points, rather than the current mm-scale distances. Laser ablation would also potentially be better in controlling the depth of the tissue evaporation, thus reducing the effect of the simultaneous evaporation of the tissues in the z-direction that has a negative effect on the classification results and annotation certainty, as hypothesized in this study.

Conclusions

We have shown that DMS-driven tissue imaging is possible with moderate resolution using an animal model. This is a novel discovery in a field dominated by MSI-driven solutions and is a step towards clinical applications. The technology could be used as a supportive method to guide pathological analysis. However, the limitations of this study need to be resolved before a commercial or independently usable research device would be feasible. In addition, the capability of detecting pathological tissue and the accuracy of tumour margin detection needs to be investigated further.

Conflict of Interest statement

Niku Oksala, Antti Roine, Markus Karjalainen, Osmo Anttalainen, Pekka Kumpulainen, and Anton Kontunen, are shareholders of Olfactomics Oy, which is about to commercialize proprietary technology for the detection of diseases by ion mobility spectrometry. The remaining authors have no conflicts of interest to declare.

This study was financially supported (or partly supported) by the Doctoral School of Tampere University; by The Finnish Foundation for Technology Promotion; by the Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital (9s045, 151B03, 9T044, 9U042, 150618, 9V044 and 9X040, 9AA057); by Competitive funding to strengthen university research profiles funded by Academy of Finland, decision number 292477; and by Tampereen Tuberkuloosisäätiö (Tampere Tuberculosis Foundation).

References

Anttalainen, O., Puton, J., Peräkorpi, K., Budzyńska, E., Eiceman, G., Sillanpää, M., 2018. Differential mobility spectrometers with tuneable separation voltage – Theoretical models and experimental findings. TrAC Trends Anal. Chem. 105, 413–423. https://doi.org/https://doi.org/10.1016/j.trac.2018.05.018

Balog, J., Sasi-Szabó, L., Kinross, J., Lewis, M.R., Muirhead, L.J., Veselkov, K., Mirnezami, R., Dezso, B., Damjanovich, L., Darzi, A., Nicholson, J.K., Takáts, Z., 2013. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. Sci. Transl. Med. 5, 1–13. https://doi.org/10.1126/scitranslmed.3005623

Chouinard, C.D., Wei, M.S., Beekman, C.R., Kemperman, R.H.J., Yost, R.A., 2016. Ion Mobility in Clinical Analysis: Current Progress and Future Perspectives. Clin. Chem. 62, 124–133. https://doi.org/10.1373/clinchem.2015.238840

Covington, J. a., van. der Schee, M.P., Edge, a. S.L., Boyle, B., Savage, R.S., Arasaradnam, R.P., 2015. The application of FAIMS gas analysis in medical diagnostics. Analyst 140, 6775–6781. https://doi.org/10.1039/C5AN00868A

Esbona, K., Li, Z., Wilke, L.G., 2012. Intraoperative Imprint Cytology and Frozen Section Pathology for Margin Assessment in Breast Conservation Surgery: A Systematic Review. Ann. Surg. Oncol. 19, 3236–

3245. https://doi.org/10.1245/s10434-012-2492-2

Golf, O., Strittmatter, N., Karancsi, T., Pringle, S.D., Speller, A.V.M., Mroz, A., Kinross, J.M., Abbassi-Ghadi, N., Jones, E.A., Takats, Z., 2015. Rapid evaporative ionization mass spectrometry imaging platform for direct mapping from bulk tissue and bacterial growth media. Anal. Chem. 87, 2527–2534. https://doi.org/10.1021/ac5046752

Jin, X., Han, J., 2010. K-Means Clustering BT Encyclopedia of Machine Learning, in: Sammut, C., Webb, G.I. (Eds.), Encyclopedia of Machine Learning. Springer US, Boston, MA, pp. 563–564. https://doi.org/10.1007/978-0-387-30164-8_425

Kabir, K.M.M., Donald, W.A., 2017. Microscale differential ion mobility spectrometry for field deployable chemical analysis. TrAC Trends Anal. Chem. 97, 399–427. https://doi.org/https://doi.org/10.1016/j.trac.2017.10.011

Komura, D., Ishikawa, S., 2018. Machine Learning Methods for Histopathological Image Analysis. Comput. Struct. Biotechnol. J. 16, 34–42. https://doi.org/https://doi.org/10.1016/j.csbj.2018.01.001

Kontunen, A., Karjalainen, M., Lekkala, J., Roine, A., Oksala, N., 2018. Tissue Identification in a Porcine Model by Differential Ion Mobility Spectrometry Analysis of Surgical Smoke. Ann. Biomed. Eng. 46, 1091–1100. https://doi.org/10.1007/s10439-018-2035-5

Nature, 2017. Histopathology is ripe for automation. Nat. Biomed. Eng. 1, 925. https://doi.org/10.1038/s41551-017-0179-5

Nazarov, E., Coy, S., Krylov, E., Miller, R., Eiceman, G., 2006. Pressure Effects in Differential Mobility Spectrometry. Anal. Chem. 78, 7697–7706. https://doi.org/10.1021/ac061092z

Orosco, R.K., Tapia, V.J., Califano, J.A., Clary, B., Cohen, E.E.W., Kane, C., Lippman, S.M., Messer, K., Molinolo, A., Murphy, J.D., Pang, J., Sacco, A., Tringale, K.R., Wallace, A., Nguyen, Q.T., 2018. Positive Surgical Margins in the 10 Most Common Solid Cancers. Sci. Rep. 8, 1–9. https://doi.org/10.1038/s41598-018-23403-5

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P.,

Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Phelps, D.L., Balog, J., Gildea, L.F., Bodai, Z., Savage, A., El-Bahrawy, M.A., Speller, A. V, Rosini, F., Kudo, H., McKenzie, J.S., Brown, R., Takáts, Z., Ghaem-Maghami, S., 2018. The surgical intelligent knife distinguishes normal, borderline and malignant gynaecological tissues using rapid evaporative ionisation mass spectrometry (REIMS). Br. J. Cancer 118, 1349–1358. https://doi.org/10.1038/s41416-018-0048-3

St John, E.R., Balog, J., McKenzie, J.S., Rossi, M., Covington, A., Muirhead, L., Bodai, Z., Rosini, F., Speller, A.V.M., Shousha, S., Ramakrishnan, R., Darzi, A., Takats, Z., Leff, D.R., 2017. Rapid evaporative ionisation mass spectrometry of electrosurgical vapours for the identification of breast pathology: Towards an intelligent knife for breast cancer surgery. Breast Cancer Res. 19, 1–14. https://doi.org/10.1186/s13058-017-0845-2

Sutinen, M., Kontunen, A., Karjalainen, M., Kiiski, J., Hannus, J., Tolonen, T., Roine, A., Oksala, N., 2019. Identification of breast tumors from diathermy smoke by differential ion mobility spectrometry. Eur. J. Surg. Oncol. 45, 141–146. https://doi.org/10.1016/j.ejso.2018.09.005

Tucker, F.L., 2012. Imaging-assisted large-format breast pathology: program rationale and development in a nonprofit health system in the United States. Int. J. Breast Cancer 2012, 171792. https://doi.org/10.1155/2012/171792

Veselkov, K.A., Mirnezami, R., Strittmatter, N., Goldin, R.D., Kinross, J., Speller, A.V.M., Abramov, T., Jones, E.A., Darzi, A., Holmes, E., Nicholson, J.K., Takats, Z., 2014. Chemo-informatic strategy for imaging mass spectrometry-based hyperspectral profiling of lipid signatures in colorectal cancer. Proc. Natl. Acad. Sci. 111, 1216–1221. https://doi.org/10.1073/pnas.1310524111

Wang, S., Tang, J., Liu, H., 2017. Feature Selection BT, in: Sammut, C., Webb, G.I. (Eds.), Encyclopedia of Machine Learning and Data Mining. Springer US, Boston, MA, pp. 503–511. https://doi.org/10.1007/978-1-4899-7687-1_101