iDMS - breakthrough in molecular imaging

Antti Vehkaoja,[1*] Anna Anttalainen,[1,2] Artturi Vuorinen,[1] Maiju Lepomäki,[1] Jorre Van Dijck[3], Anton Kontunen,[1,2] Meri Mäkelä,[1,2] Markus Karjalainen,[1,2] Niku Oksala,[1,2] Antti Roine[2]

[1]Faculty of Medicine and Health Technology, Tampere University, Arvo Ylpön katu 34, 33520, Tampere, Finland; [2]Olfactomics Ltd., Korkeakoulunkatu 7, Kampusareena, FI-33720 Tampere, Finland; [3]Hasselt University, Hasselt, Belgium.
*Corresponding author: antti.vehkaoja@tuni.fi

ABSTRACT
iDMS project has developed technology for molecular imaging focusing on the analysis of tissues and especially cancer and related specific molecules. We have for example successfully demonstrated quantitative estimation of the concentration of lecithin that can be considered as a representative biomarker of cancerous tissue for the first time with differential mobility spectrometry technology and convolutional neural networks. In addition, we demonstrated 94% accuracy in cancer tissue identification with the developed laser evaporation platform. The end-users of our technology include pathologists and oncologists but applications are found also in other areas of life sciences.

Keywords: Cancer tissue identification, differential mobility spectrometry, laser evaporation, machine learning analytics

1. INTRODUCTION

Molecular imaging is an emerging field that has several important application areas, one which is the improvement of the efficiency of pathological examinations. Pathology is currently an extremely labour-intensive area as the examinations are performed visually. Pathological examinations are also time consuming due to complex preparation of the samples. In many cases e.g. during surgery obtaining analysis result from the removed tissue sample immediately would be valuable but is limited due to high costs of special preparation techniques.
Current technology used in molecular imaging relies on mass-spectrometry, which is the gold standard in material composition analysis. It is however complex and costly technology and therefore its availability is limited and generalization severely hindered. Differential mobility spectrometry (DMS) is an alternative technology that is significantly simpler and therefore cheaper to manufacture and would enable bringing the cost of molecular imaging equipment down by an order of magnitude. In iDMS project we have:

1. Developed a cost-effective, laser based tissue imaging system utilizing differential mobility spectrometry
2. Produced a dataset of > 3500 animal tissue samples and a pilot dataset with human tumor tissue
3. Created a data analysis pipeline utilizing conventional and deep learning methods developed for tissue classification with > 90% accuracy.
4. Performed quantitative analysis of DMS data for the first time: Phospholipid concentration estimation using linear and state-of-the-art non-linear regression methods.

2. STATE OF THE ART

Histopathological examination of tumor specimens is the technique currently used in clinical practise to achieve precise diagnosis and determine surgical margin status after surgery. During margin examination the specimen is painted with insoluble dye and sectioned perpendicular to the inked surface. Due to the sampling technique, the examination does not allow comprehensive analysis of the entire specimen and is time-consuming. [1]
The state of the art in molecular imaging has been the utilization of mass spectrometry -based technology. The most common among these techniques are matrix assisted laser desorption/ionization (MALDI), secondary ion mass spectrometry (SIMS), desorption electrospray ionization (DESI), and rapid evaporative ionization mass spectrometry (REIMS). [2-4] The weakness of mass spectrometry imaging techniques are their complexity and economic strain which impair their feasibility outside research facilities.
An efficient tissue imaging method in which the composition of the surgical specimen could be analysed prior to its histopathological examination would support the work of a pathologist and enhance sampling. The method could be beneficial particularly in the selection of surgical margin samples to ensure complete carcinoma removal.

3. BREAKTHROUGH CHARACTER OF THE PROJECT

The major breakthrough of the iDMS project is the developed laser-DMS molecular imaging methodology that is demonstrated during in several application cases. The DMS-based molecular and tissue imaging system developed in this project overcomes many of the challenges that are faced with the MS-based methods and in gold standard histopathological examination, while retaining an adequate accuracy and speed of analysis. In practice, this means that iDMS has great potential of saving valuable healthcare resources by complementing or even replacing the existing pathological examination methods. A comparison between the three methodologies and their advantages/disadvantages is shown in Tab. 1.

Tab. 1. Comparison of methods for pathological tissue imaging

Attribute	iDMS	MSI	Histopat hology
Cost	+++	-	-
Speed of analysis	++	++	-
Accuracy	+	++	+++
Robustness	+++	-	+
Ease-of-use	+++	+	+
Maintenance	++	-	+
Required resources (time & labor)	+++	+	-

4. PROJECT RESULTS

Laser evaporation -based volatile compound analysis system

We have developed a laser evaporation -based sampling system for tissue classification and analysis. The system consists of a pulsed 10.6 μm carbon dioxide laser with a moving mirror xy-gantry assembly used to direct the laser beam. The laser system is attached to a DMS analyser through a sample gas filtering and dilution stage. The modules of the system are controlled with two Raspberry Pi computers through a graphical user interface. Fig. 1 illustrates the units of the system. Fig. 2 shows the laser collimation lens and the sample vapour collection nozzle that forms an air bed around the evaporation area thus preventing external undesired volatile compounds from being transported to the analyser.

Porcine tissue identification

A dataset consisting of tissue samples from 18 pigs was collected with the developed system during a period of five months. The tissues were: skeletal muscle, adipose tissue, breast tissue, liver, and bone tissue. Altogether 3561 evaporated tissue samples were used in training machine learning models for tissue classification and an additional external validation set of 185 samples for evaluation of the system performance. The evaluated machine learning approaches included shrinkage linear discriminant analysis (sLDA), support vector machine (SVM), and convolutional neural network (CNN) model. The best overall accuracies of 86.4% and 81.6% were obtained with the CNN method for the leave one measurement day out and external validation data set, respectively. The other methods provided 5-10 percentage points lower accuracies. Table 2 presents the confusion matrix for CNN model in leave-one-day-out cross validation. True class values are on each column. [5]

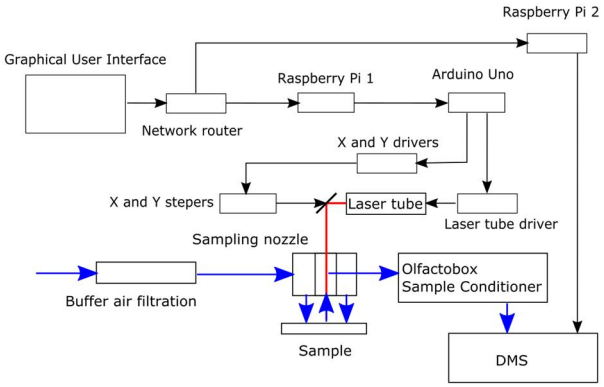

Fig. 1. Block diagram of the developed tissue analysis system.

Tab. 2. Confusion matrix of leave one day out cross validation results for porcine tissue identification test.

	Muscle	Fat	Breast	Liver	Bone
Muscle	***924***	*37*	*20*	*6*	*15*
Fat	*42*	***1223***	*16*	*0*	*72*
Breast	*42*	*39*	***200***	*0*	*79*
Liver	*1*	*0*	*0*	***215***	*1*
Bone	*21*	*30*	*22*	*43*	***513***

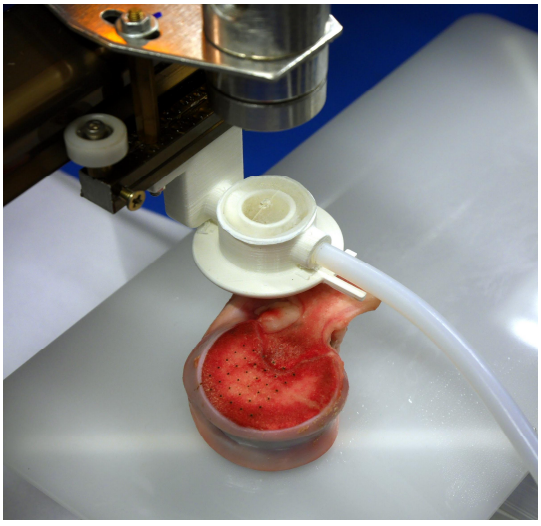

Fig. 2. Porcine femur under measurement in the laser sampling system.

Breast cancer pilot

Three breast cancer tumor samples were analysed for piloting the ability of the laser-based system to distinguish between benign and cancer tissue samples. The samples provided altogether 260 laser incisions. analysis spots. The samples were annotated by a pathologist both macroscopically from images and microscopically from histopathological samples. Due to the limited size of the dataset, random 10-fold cross validation analysis approach was chosen to evaluate robustness of the results.. 94% accuracy was obtained with SVM classifier using radial kernel. Figure 3 shows examples of DMS dispersion matrices obtained from breast cancer and healthy tissue. Areas in the spectra that mostly differentiate the malignant and benign tissue are marked in the spectra. Figure 4 illustrates the analysed breast cancer samples with an overlay of the classification results. [5]

Quantitative estimation of lecithin concentration

Motivation behind studying the estimation of lecithin concentration was that mass spectrometry studies of surgical smoke have shown that phospholipid content and concentration are the main differentiating factor between benign and malignant tissues. Lecithin is a mix of phospholipids that are also found in tissues. In this study the samples were collected from the headspace of a vessel containing lechitin solution in various concentrations.

The studied concentrations varied between 0-6.2 mg mL^{-1}. Altogether 960 samples in 9 different concentrations of lecithin and water were measured and used to train various regression models: sparse linear regression with ridge, lasso and elastic net regularization methods as well as partial least squares (PLS) and CNN regression models. 10-fold cross-validation (CV) external 100-sample dataset were used in evaluation.

The best cross-validation results were obtained with CNN model with RMSE = 0.38 mg mL^{-1}. For the external validation set, the best results were obtained with regularized linear regression methods. RMSE = 0.40 mg mL^{-1} - 0.41 mg mL^{-1}. The reason for the better success of the less complex methods is likely in the relatively small size of the dataset. The results demonstrate that DMS technique is sufficiently sensitive to detect biologically relevant changes in the phospholipid concentration. The use of neural networks in concentration prediction from DMS measurements also lays a foundation for its wider analytical usage. [6]

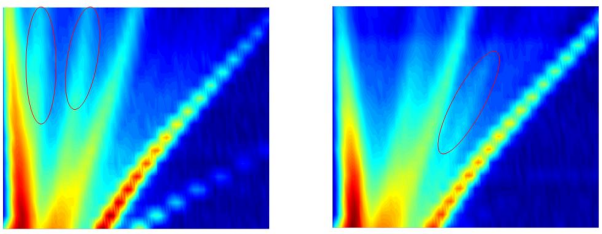

Fig. 3. DMS spectra of healthy (right) and cancerous tissue (left). The main differentiating areas are highlighted.

Antibiotic detection in tissue matrix

We made a pilot study to detect antibiotic contaminants from food. We used homogenized porcine muscle tissue spiked with 10000 µg/g of penicillin G (Fig. 5). The concentration used in this preliminary test is approximately 1000 fold higher than physiological 10 µg/g concentration. The pilot produced 91.67% positive predictive value with 150 measured samples and using a simple LDA classifier. Reduced antibiotic concentrations in combination with significantly larger number of samples will be next measured for being able to use CNN based classifiers and evaluate the applicability of the method.

5. FUTURE PROJECT VISION

After a successful execution of Attract Phase 1 project and encouraged by the promising result obtained, iDMS team is committed to continue the transformation of this technology from a demonstrator to a scalable solution that is ready for commercialization.

5.1. Technology Scaling

In Attract Phase 1 we have reached the technology readiness levels (TRL) of 3 and 4 in both quantitative estimation of analyte concentration in the sample and in qualitative classification of tissue types, respectively. Our next target is to demonstrate and validate the technology and its benefits in different phases of the treatment of various cancers thus raising the maturity of the technology to the critical level where it can draw attention of investors and future end users (TRL level 6). In addition to cancer treatment, the technology is highly applicable to various other uses in biological material characterization and e.g. in disease diagnosis.

We have identified critical factors needed for successful scaling of the research results to commercially viable solutions. These include:

- Robust technology
- Scalable, high-volume applications
- Complementary core team
- Strong collaboration network

Thus, the goal of iDMS Attract Phase 2 project is to transform the highly promising proof-of-concept to a technology platform that combines

1) Proven technology
2) Several validated applications
3) Active research use base
4) A team ready for commercialization

In order to enable cost-effective molecular imaging for multiple applications in the area of life sciences.

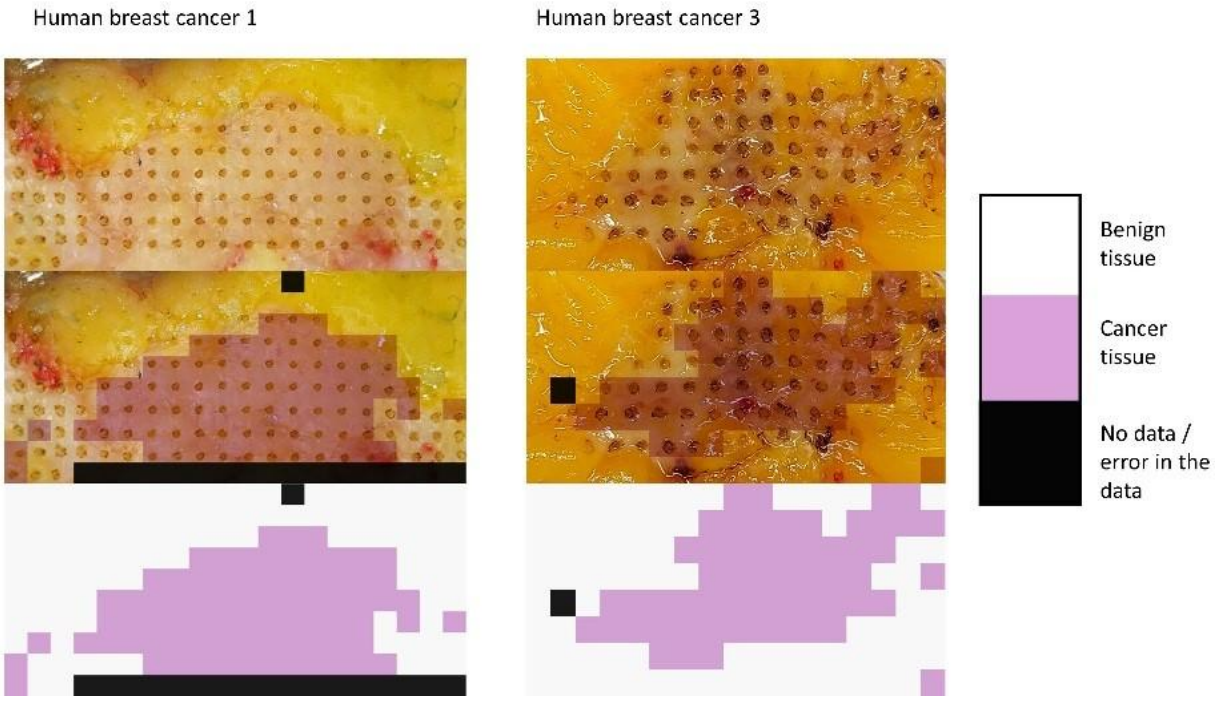

Fig. 5. Visualization of results for the macroscopically annotated breast cancer samples with SVM (radial kernel).

5.2. Project Synergies and Outreach

In Attract Phase 2 we seek to extend our consortium by collaborating with critical technology providers as well as with end users and utilizers of the technology representing various fields of cancer medicine. We also welcome representatives of new utilization areas who see large potential and benefits in this technology.
Providers of critical technology include laser technology developers and manufacturers. Our long term vision is to combine the iDMS analysis technology with extended reality technology to create applications that can substantially improve the workflow of pathologists and also surgeons. Therefore, we also seek for partners focusing on the development of XR systems and their applications.

5.3. Technology application and demonstration cases

In Attract phase 2, we will have cancer diagnostic and treatment as the main application area. Different types of cancers require significantly different diagnostics approaches and therefore can benefit from the possibilities offered by DMS technology in various ways. We will perform parallel exploration of high-potential applications in collaboration with a netrowork of leading international academic partners to secure the highest level of domain-knowledge. The envisioned focus applications include:

- Quantitative estimation of HER-protein concentration for breast cancer treatment planning
- Detection of specific markers describing IDH-mutations in brain tumor assessment to assist in treatment
- Automated Gleason scoring method for prostate cancer grading

Focusing on a few selected applications will secure the limited resources being focused in a right way so that it is possible to develop the solutions in a commercialization-ready state during the project.

Additionally, we will encourage close collaboration of different users of the technology by creating a curated repository where new control algorithms, anonymized datasets and other features can be shared by the researchers. This way, individual teams will gain access to a large body of work and researchers remain motivated to share their work with the community.

5.4. Technology commercialization

Besides validating the iDMS technology in the selected specific applications named in the previous section, the iDMS consortium will also prepare commercialization of the technology with following actions. We will:

1. Evaluate possible business models for their suitability for commercializing such a high tech solution.

2. We will complement our team with required new competences that have not been relevant in the preliminary research phase. These include for example sales and business development expert.
3. We will prepare investor materials and approach potential medical technology investors.
4. We will protect the critical value adding intellectual properties generated in the development projects.

5.5. Envisioned risks

The applications developed in the Attract Phase 2 project are completely new. Therefore, despite the extensive background research done for selecting these specific applications, there are risks of discovery of unexpected obstacles along the way of the research and development. Such obstacles may include:

Risk: Inadequate performance. The variation of the VOC content may be too small for discriminating cancerous and healthy tissue with adequate performance in cases where the amount of cancer cells in the sample is small.

Mitigation: The resolution of the created DMS dispersion plot may be increased but in this case the measurement time of a single sample is increased. The use of label molecules is also a possibility but this will reduce the strength of the proposed method, i.e. straightforward sampling and operation.

Risk: Minute total benefits. The overall added value of the proposed new method shows to be lower than expected in a specific application.

Mitigation: The proposed applications are carefully chosen and represent the most prevalent types of cancer in both men and women. However, only thorough evaluation performed during the Attract Phase 2 will reveal the economical sustainability of the solutions. I may show that some other phase of the treatment process may be more feasible and can in that case also be realized.

Due to the aforementioned risks the development is continuously critically monitored to find out in an early state if any of the chosen application presents fundamental difficulties and the team is in such case ready for pivoting and either choosing an alternative approach or focusing the development efforts to the remaining applications.

5.6. Liaison with Student Teams and Socio-Economic Study

In Attract Phase 2 iDMS team will arrange at least two full day workshops with student teams to co-create new use cases for the technology. iDMS team will also provide trainee positions for students as we have already done in Attract Phase 1.

iDMS team is also happy to contribute to the socio-economic study being conducted as a part of Attract Phase 2 project by providing interviews and sharing our expertise on the impact of medical technology especially in the cancer treatment area.

6. ACKNOWLEDGEMENT

iDMS project team wishes to thank Pathologist Teemu Tolonen, FIMLAB Laboratories for providing the the reference annotations for the breast cancer pilot.

This project has received funding from the ATTRACT project funded by the EC under Grant Agreement 777222.

7. REFERENCES

[1] Morrow, M., 2009. Breast conservation and negative margins: How much is enough?, The Breast. 18: pp. S84-S86.

[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, Analytical Chemistry, 87: pp. 2527-2534.

[3] Veselkova, K.A., Mirnezamib, R., Strittmattera, N., Goldinc, R.D., Kinrossb, J., Spellerc, A.V.M., Abramovd, T., Jonesa, E.A., Darzib, A., Holmesa, E., Nicholsona, J.K., Takatsa, Z., 2014. Chemo-informatic strategy for imaging massspectrometry-based hyperspectral profiling of lipid signatures in colorectal cancer, Proceedings of the National Academy of Sciences, 111(3): pp. 1216-1221.

[4] Mirnezamia, R., Spagoub, K., Vorkasb, P.A., Lewisb, M.R., Kinrossa, J., Wantb, E., Shionc, H., Goldind, R.D., Darzia, A., Takatsb, Z., Holmesb, E., Cloarecb, E., Nicholsonb, J.K., 2013. Chemical mapping of the colorectal cancer microenvironment via MALDI imaging mass spectrometry (MALDI-MSI) reveals novel cancer-associated field effects, Molecular Oncology, 8(1): pp. 39-49.

[5] Lepomäki, M., Anttalainen A., Vuorinen A., Tolonen T., Kontunen A., Karjalainen M., Vehkaoja A., Roine A., Oksala N., 2020, Tissue imaging with Differential Ion Mobility Spectrometry (DMS) and laser sampling. Unpublished manuscript.

[6] Anttalainen, A., Mäkelä, M, Kumpulainen, P., Vehkaoja, A., Anttalainen, O., Oksala, N., Roine, A., 2020 Predicting lecithin concentration from differential mobility spectrometry measurements with linear regression models. Unpublished manuscript.