Tampere University

ANTON KONTUNEN

Tissue Identification by Differential Mobility Spectrometry

Tampere University Dissertations 557

Tampere University Dissertations 557

ANTON KONTUNEN

Tissue Identification by Differential Mobility Spectrometry

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Medicine and Health Technology
of Tampere University,
for public discussion in the auditorium S2
of the Sähkötalo, Korkeakoulunkatu 3, Tampere,
on 11 March 2022, at 9 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Medicine and Health Technology
Finland

Responsible supervisor and Custos	Associate Professor (tenure track) Antti Vehkaoja Tampere University Finland	
Supervisors	MD, PhD Antti Roine Olfactomics Oy Finland	Professor Niku Oksala Tampere University Finland
Pre-examiners	PhD Marjaana Nousiainen Finnish Food Authority Finland	Assistant Professor Heikki Nieminen Aalto University Finland
Opponent	Associate Professor Brendan Kennedy The University of Western Australia Australia	

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Cover design: Roihu Inc.

ISBN 978-952-03-2302-8 (print)
ISBN 978-952-03-2303-5 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-2303-5

PunaMusta Oy – Yliopistopaino
Joensuu 2022

PREFACE

In the preface of my master's thesis in 2017, I stated that I could not have hoped for a better or more interesting thesis topic. This opinion still holds true after continuing the project for more than four years. I am extremely appreciative for the opportunity to work on this research with some of the most talented and motivated people on the planet.

This thesis project has been a demanding undertaking, which I hoped I would have been able to finish a bit sooner. Nevertheless, the things I have learned during this project are invaluable and the personal development could not have been possible if I had not continued my academic career this far. As this thesis project went hand-in-hand with our efforts to commercialise the idea, almost every day was filled with new interesting and exciting developments, both in terms of academic research, and medical device development in practice. Regardless of what happens in the future, I am forever grateful for this experience.

The research presented in this thesis started out in Tampere University of Technology and continued in its successor institute, Tampere University. I am grateful for the opportunity to work towards my graduation with the funding from the Tampere University Doctoral school. I am also grateful for the personal funding I received from the Finnish Cultural Foundation South Savo regional fund, The Finnish Foundation for Technology Promotion, and Emil Aaltonen Foundation.

I am greatly indebted to the pre-examiners of this thesis, PhD Marjaana Nousiainen, and Asst. Prof. (tenure track) Heikki Nieminen for their contribution and insightful comments. I am also grateful to Assoc. Prof. Brendan Kennedy, who agreed to act as my opponent in my doctoral defence. I also want to thank my three supervisors, Assoc. Prof. (tenure track), D.Sc., Antti Vehkaoja; MD, PhD, CEO, Antti Roine; and Prof., MD, PhD, D.Sc., Niku Oksala, who were integral for the completion of this work. Ever since I met them, they have inspired with their ideas and experience, and motivated me to work ever harder and smarter. They have invested a lot of their time and effort in guiding me with the thesis project. Their productiveness, responsiveness, and ingenuity never cease to amaze me. I also want

to express my gratitude towards my steering group member D.Sc. Ville Rantanen, who worked as my supervisor before my doctoral studies and who has always been willing to offer his help and expertise, even after moving on from the academic field.

I want to thank all of my co-authors. Without them, the publications that work as the basis of this thesis would not have been possible. This is especially true for Maiju Lepomäki, Jalmari Tuominen, and Ulla Karhunen-Enckell, with whom I shared first authorships with, and without whom I would have been completely lost with most of the medical aspects of the publications.

I need to thank all of my colleagues from the faculty of medicine and health technology, who I have had a genuine pleasure to work with during these years. The discussions, coffee breaks and occasional recreational activities brought joy amidst the hard work. I also want to thank my amazing colleagues from Olfactomics: Markus Karjalainen, Meri Mäkelä, Anna Anttalainen, Osmo Anttalainen, Mikko Koskenranta, Arttu Havia, Pekka Kumpulainen, and all others, who have worked with me during these years and who have always made me feel as a part of something bigger.

I also want to thank my friends in helping me keep my work-life balance in check. My siblings and parents have also always supported me in their own right, every step of the way, and I would not have gotten to this point without them, thank you. Finally, I want to thank my fiancée Kristiina. You are the light of my life.

"*Don't be guided by fear of failure*" – Ville-Petri Friman

Tampere, January 2022

ABSTRACT

The human genome is constantly changing due to natural mutations and environmental exposure. As these changes accumulate over our lifetime, it increases the likelihood of the creation of cells that proliferate uncontrollably and ultimately invade surrounding tissue and the blood circulation or the lymphatic system. This type of malignant neoplasm, more commonly known as cancer, is a disease that either directly or indirectly affects the majority of the population as one of the leading causes of death. Cancer is a versatile disease that can affect practically any part of the body. Depending on the tissue of origin and the aggressiveness of the malignancy, the treatment options, prognosis and mortality rates can vary significantly. In general, the role of cancer as a cause of death is constantly increasing, and despite significant global financial investments and decades of research, new and better methods of treatment and diagnosis are in continuous demand.

One particular area that requires more attention and innovation is the surgical treatment of solid cancers. The general aim of surgical treatment is to remove all malignant cells from the patient's body – that is to say, to achieve a negative surgical margin. The resected tumour has a negative margin, when the outermost surface area has no cancerous cells. However, in a considerable number of surgeries, the removal is incomplete. The resulting residual cancer almost always triggers additional treatment steps, which often involve a reoperation. The need for a reoperation is a major detriment for the well-being of the patient, and the added healthcare costs are substantial. If the number of avoidable reoperations could be halved from their current level, the saving potential in annual global healthcare costs would already be measured in billions of dollars.

The reason why the problem of reoperations persists despite the notable financial incentives lies in the difficulty of discriminating malignant tissue from benign, especially during a surgical procedure. The molecular contents that define the structure and function of a cell are different depending on the organ of origin, and similar differences are also present between malignant and benign cells. The biomolecules that enable the identification of the types of tissues are called biomarkers, and the research on this area has revealed hundreds of proteins, fatty

acids and metabolic products that exhibit differences in quantities based on tissue malignancy. However, the variation of specific marker molecules is often high, and the molecular differences rarely translate into clear macroscopic differences. This means that visual assessment of the margin between benign and cancerous tissue is extremely challenging. Still, almost all surgeons rely only on visual assessment and palpation in cancer surgeries. The challenge of complete excision is further accentuated by the current resection guidelines that instruct surgeons to preserve as much non-cancerous tissue as possible. This aim and its subjective execution lead not only to high variation in positive margin rates between institutions and regions, but also to a high number of required reoperations in general. To reduce the reoperations caused by positive surgical margins, several technologies have been studied and introduced to aid in intraoperative tissue identification, but the clinical adoption has been limited due to various impeding factors involved in their use.

In this thesis, a concept that could potentially be used in the assessment of the intraoperative surgical margin is introduced through five scientific publications that concentrate on the evolution and feasibility of the technology in tissue identification. The basis of the technology is the measurement of surgical smoke with differential mobility spectrometry (DMS). DMS is a measurement technology that provides information on the molecular content of a gaseous sample in atmospheric pressure by means of ionisation and subsequent differentiation of the ions in a high-strength asymmetric electric field. DMS is comparable to mass spectrometry (MS), and even though the analytical performance of MS is better, the reduced complexity, smaller size and lower cost of DMS make it an advantageous option. DMS has been used as a standalone measurement instrument in many types of general gas measurement applications and in some biomedical applications, such as breath analysis, but the context of use has always permitted a controlled environment and a relatively long measurement duration. Thus, the real-time application of surgical smoke measurement requires additional hardware and parameter optimisation. In addition, raw DMS measurement data do not provide directly quantifiable information on certain biomolecules, but rather a comprehensive spectrum of all contents in the sample combined. This means that the interpretation and identification of tissue type from the DMS output spectra is not trivial and involves a high number of dimensions that are most effectively analysed by means of machine learning. The interdisciplinary aspects of the system and their combined function and performance in tissue identification are the focus of this thesis.

In the first three publications included in the thesis, the focus was on studying the overall feasibility of tissue identification and its possibilities with animal tissues and clinically relevant breast cancer samples. The results in laboratory conditions with controlled sampling were promising, and the diagnostic performance demonstrated the potential of the technology in tissue identification. In Publication IV, the system was modified to accommodate real-time measurements and to relay the classification information immediately after the measurement. The results demonstrated the feasibility of real-time tissue identification with the system, albeit in laboratory conditions and in a porcine model. In the final study, a prototype system was used intraoperatively during breast cancer surgeries. The results of this study were not comparable to the laboratory results in respect to diagnostic performance but indicated that the system can be adapted to the surgical workflow with minimal intrusiveness to provide information on the operated tissue.

Overall, the results of this study indicate that a DMS-based tissue identification system has the potential to be used in real-time applications to identify tissue types with adequate diagnostic performance. With further development, the system presented in this thesis could fulfil the need for a surgical margin assessment device that would reduce avoidable reoperations of solid cancers and thus protect the well-being of cancer patients.

TIIVISTELMÄ

Perimämme muuttuu jatkuvasti luonnollisten mutaatioiden sekä ulkoisten tekijöiden vaikutuksesta. Muutosten kumuloituessa elämämme aikana hallitsemattomasti jakaantuvien ja ympäröiviin kudoksiin sekä lymfaattiseen järjestelmään ja verenkiertoon tunkeutuvien solujen syntymisen todennäköisyys kasvaa. Tämän tyyppistä pahanlaatuista solukasvua kutsutaan syöväksi. Syöpä vaikuttaa joko suoraan tai epäsuorasti suurimpaan osaan ihmisistä yhtenä yleisimmistä kuolinsyistä. Syöpä on monimuotoinen tauti, joka voi syntyä käytännössä mihin kehon osaan tahansa. Riippuen syövän kohdekudoksesta ja kasvun aggressiivisuudesta mahdolliset hoitomuodot, selviytymisennusteet ja kuolleisuus vaihtelevat huomattavasti. Yleisesti syövän rooli kuolinsyynä kuitenkin korostuu jatkuvasti, ja huomattavasta rahallisesta panostuksesta ja vuosikymmenten tutkimustyöstä huolimatta uusille ja paremmille hoito- ja diagnosointimenetelmille on jatkuva tarve.

Kiinteiden syöpien leikkaushoito on yksi erityisalue, joka hyötyisi uusista, hoitoa tehostavista innovaatioista. Syövän leikkaushoidossa on yleisesti tavoitteena poistaa kasvain elimistöstä täydellisesti ja täten saavuttaa negatiivinen tervekudosmarginaali. Huomattavassa osassa syöpäleikkauksia poisto on kuitenkin epätäydellinen. Tällöin potilaaseen jääneet syöpäsolut vaativat jatkohoitotoimenpiteitä, joihin yleensä sisältyy myös syövän uusintaleikkaus. Uusintaleikkauksen tarve on erittäin vahingollista potilaan yleiselle hyvinvoinnille ja tuo mukanaan huomattavia lisäterveydenhuoltokustannuksia. Jos vältettävissä olevien uusintaleikkausten määrä voitaisiin puolittaa nykyisestä, säästöjä mitattaisiin jo miljardeissa.

Selkeästä säästöpotentiaalista huolimatta syöpien turhat uusintaleikkaukset ovat edelleen ratkaisematon ongelma johtuen etenkin leikkauksenaikaisista haasteista erottaa hyvänlaatuinen kudos pahanlaatuisesta. Solujen rakenteen ja toiminnan määräävä molekulaarinen sisältö eroaa riippuen solujen syntykudoksesta, ja samankaltaisia eroja havaitaan myös pahanlaatuisten ja hyvänlaatuisten solujen välillä. Biomolekyylejä, jotka mahdollistavat kudostyyppien erojen havaitsemisen, kutsutaan biomarkkereiksi tai bioilmaisimiksi, ja tutkimuksissa onkin löydetty satoja proteiineja, rasva-aineita ja aineenvaihduntatuotteita, joiden pitoisuus solussa vaihtelee hyvänlaatuisen ja pahanlaatuisen kudoksen välillä. Tiettyjen

biomarkkereiden pitoisuuksien vaihteluvälit hyvänlaatuisissa ja pahanlaatuisissa kudoksissa ovat kuitenkin erittäin suuria, ja molekyylitason erot kudosten välillä aiheuttavat harvoin selkeää makroskooppisesti näkyvää muutosta. Siksi syöpäkudoksen ja tervekudoksen välisen rajan silmämääräinen arvioiminen on erittäin haastavaa. Silti lähes kaikki syöpäkirurgit käyttävät ainoastaan visuaalista arviointia ja tunnustelua operaatiotilanteessa. Lisää haasteellisuutta syövän kokonaispoistoon tuovat myös nykysuositukset, joiden mukaan syövän ympäriltä poistetun tervekudoksen määrä pyritään minimoimaan. Tämä tavoite ja rajan subjektiivinen arviointi johtavat suureen hajontaan eri maiden ja sairaaloiden positiivisten marginaalien määrissä sekä yleisesti korkeaan uusintaleikkausten määrään. Positiivisista tervekudosmarginaalilöydöksistä johtuvia uusintaleikkauksia on pyritty vähentämään tutkimalla ja ottamalla käyttöön useita erilaisia leikkauksenaikaista kudostunnistusta auttavia menetelmiä, mutta niiden kliininen käyttö on ollut rajallista johtuen kunkin menetelmän rajoitteista ja haitoista.

Tässä väitöskirjassa esitellään kudostunnistusjärjestelmä, jota voidaan mahdollisesti tulevaisuudessa hyödyntää leikkauksenaikaisessa tervekudosmarginaalin arvioinnissa. Järjestelmän kehitystä ja soveltuvuutta kudostunnistukseen tarkastellaan viiden osatyön kautta. Järjestelmä pohjautuu sähkökirurgiassa tuotetun kudossavun mittaamiseen liikkuvuuserospektrometrialla (*differential mobility spectrometry*, DMS). DMS on normaali-ilmanpaineessa toimiva mittausteknologia, joka tuottaa informaatiota kaasumaisen näytteen molekulaarisesta rakenteesta erottamalla ionisoidut molekyylit toisistaan voimakkaassa, epäsymmetrisesti muuttuvassa sähkökentässä. DMS vertautuu massaspektrometriaan (MS) mutta on analyyttiseltä suorituskyvyltään sitä heikompi. DMS-teknologian etuna on kuitenkin sen yksinkertaisuus, pienempi koko sekä pienemmät kustannukset MS-teknologiaan verrattuna. DMS-teknologiaa on aiemmin käytetty itsenäisenä mittausmenetelmänä erilaisissa kaasumittaussovelluksissa sekä biolääketieteellisessä käytössä muun muassa hengitysilman mittaamiseen. Nämä sovellukset ovat kuitenkin aina sallineet kontrolloidun ympäristön ja suhteellisen pitkän mittauksen keston. Siksi reaaliaikainen DMS-pohjainen sovellus vaatii ympärilleen lisälaitteistoa ja järjestelmän parametrien optimointia. Lisäksi DMS-data ei suoraan tuota määrällistä tietoa näytteessä olevista biomolekyyleista vaan luo pikemminkin kokonaiskuvan näytteen sisältämien aineiden seoksesta. Spektrin tulkinta ja kudostyypin määritys ei siis ole suoraviivaista, ja yhdestä näytteestä saatavan suuren datamäärän vuoksi analysointi soveltuu parhaiten koneoppimismenetelmille. Järjestelmän poikkitieteellinen näkökulma sekä

kokonaisuuden toiminnan ja suorituskyvyn tutkiminen kudostunnistuksessa ovat tämän väitöskirjan pääsisältö.

Väitöskirjan kolmessa ensimmäisessä osatyössä tavoitteena oli tutkia menetelmän soveltuvuutta kudostunnistukseen eläinkudosnäytteillä sekä ihmisen rintasyöpänäytteillä. Tulokset laboratorio-olosuhteissa hallitulla näytteentuotolla olivat lupaavia, ja diagnostinen suorituskyky osoitti teknologian potentiaalin kudostunnistuksessa. Neljännessä osatyössä laitteistoa muokattiin mahdollistamaan reaaliaikaiset mittaukset sekä luokittelutuloksen esitys välittömästi mittauksen jälkeen. Tulokset osoittivat, että järjestelmä soveltuu reaaliaikaiseen kudostunnistukseen vähintään eläinnäytteillä laboratorio-olosuhteissa. Viidennessä osatyössä järjestelmää käytettiin rintasyöpäleikkauksissa. Diagnostisen suorituskyvyn osalta tulokset eivät olleet vertailukelpoisia laboratoriotutkimuksiin, mutta tutkimus osoitti, että järjestelmän integroiminen osaksi syöpäkirurgiaa onnistuu käyttäjiä häiritsemättä ja että se pystyy tuottamaan informaatiota leikatusta kudoksesta operaation aikana.

Kokonaisuudessaan väitöskirjatutkimuksen tulokset osoittavat DMS-pohjaisen kudostunnistusjärjestelmän potentiaalin ja soveltuvuuden reaaliaikaiseen käyttöön riittävällä diagnostisella suorituskyvyllä. Tulevaisuudessa tässä työssä esitetty järjestelmä voi jatkokehityksen jälkeen toimia syöpäkirurgin apuna tervekudosmarginaalin tunnistuksessa ja auttaa suojelemaan syöpäpotilaiden hyvinvointia vähentämällä tarpeettomia syövän uusintaleikkauksia.

CONTENTS

ABBREVIATIONS AND SYMBOLS

^{241}Am	americium-241
AUC	area under the curve
ATAS	automatic tissue analysis system
CNN	convolutional neural network
DCIS	ductal carcinoma in situ
DMS	differential mobility spectrometry
DNA	deoxyribonucleic acid
FAIMS	field asymmetric ion mobility spectrometry
FN	false negative
FP	false positive
FSA	frozen section analysis
FSFS	forward sequential feature selection
GUI	graphical user interface
IMS	ion mobility spectrometry
LDA	linear discriminant analysis
LOOCV	leave-one-out cross-validation
MS	mass spectrometry
NPV	negative predictive value
OCT	optical coherence tomography
PLA	polylactic acid
PPV	positive predictive value
REIMS	rapid evaporative ionization mass spectrometry
RIP	reactant ion peak

ROC	receiver operating characteristic
Td	Townsend
TN	true negative
TP	true positive
U_C	compensation voltage
U_{SV}	separation voltage
V_{B+}	positive bias voltage
V_{B-}	negative bias voltage
VOC	volatile organic compound

$\alpha(E/N)$	function describing the dependence of mobility on the ratio of electric field to neutral molecule number density
E	electric field strength
$H^+(H_2O)_n$	protonated water cluster
$J(\vec{w})$	criterion function
K	mobility coefficient
K_0	reduced mobility coefficient
$K(E/N)$	field-dependent mobility
M	neutral analyte molecule
m	sample mean
N	number density of neutral gas molecules ($1/m^3$)
$O_2^-(H_2O)_n$	deprotonated water cluster
P	pressure
S_B	between-class scatter matrix
S_W	within-class scatter matrix
T	temperature
v	drift velocity
$\vec{w}$	projection vector

LIST OF ORIGINAL COMMUNICATIONS

Publication I **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. Annals of biomedical engineering, 46(8), 1091–1100.

Publication II 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. European Journal of Surgical Oncology, 45(2), 141–146.

Publication III **Kontunen, A.***, Tuominen, J.*, Karjalainen, M., Anttalainen, O., Tolonen, T., Kumpulainen, P., Lepomäki, M., Vehkaoja, A., Oksala, N., & Roine, A. (2020). Differential mobility spectrometry imaging for pathological applications. Experimental and Molecular Pathology, 117, 104526.

Publication IV **Kontunen, A.**, Karjalainen, M., Anttalainen, A., Anttalainen, O., Koskenranta, M., Vehkaoja, A., Oksala, N., Roine, A. (2021). Real Time Tissue Identification from Diathermy Smoke by Differential Mobility Spectrometry. IEEE Sensors Journal, 21(1), 717–724.

Publication V **Kontunen, A.***, Karhunen-Enckell, U.*, Karjalainen, M., Anttalainen, A., Kumpulainen, P., Pitkänen, L., Anttalainen, O., Vehkaoja, A., Oksala, N., Roine, A. (2021). Tissue Identification from Surgical Smoke by Differential Mobility Spectrometry: an in vivo study. IEEE Access, 9, 168355-168367

*Equal contribution

AUTHOR'S CONTRIBUTION

The author of this thesis was the sole main author in Publications I and IV. In Publications II and III and V, the lead authorship was shared with one other person with equal contribution. The authors' contributions in each publication, with additional details on other authors in the shared publications, are described below.

Publication I **The author** planned and conducted the measurements and data analysis, assisted in the creation of the methodology, and led the manuscript writing process and submission as the corresponding author.

Publication II **The author** planned and conducted the measurements and the writing process together with Maiju Lepomäki (née Sutinen) with equal contribution. Lepomäki led the manuscript writing process and submission in the role of the corresponding author. The data analysis was carried out by the author with assistance from Lepomäki. The author assisted in the creation of the technical methodology. Maiju Lepomäki and Teemu Tolonen were in charge of the medical methodology (samples, histopathology).

Publication III **The author** planned and conducted the measurements, data analysis and the writing process in co-operation with Jalmari Tuominen, with equal contribution. The author was in charge of the supervised classification, and Tuominen was in charge of the unsupervised clustering. The author led the manuscript writing process and submission in the role of the corresponding author. The author assisted in the creation of the technical methodology.

Publication IV **The author** planned and conducted the measurements and data analysis (with Anna Anttalainen contributing to the convolutional neural network analysis), assisted in the creation of the methodology, and led the manuscript writing process and submission as the corresponding author.

Publication V **The author** planned and conducted the measurements and the writing process together with Ulla Karhunen-Enckell, with equal contribution. The author oversaw the function of the system in the operating room, and Karhunen-Enckell performed the operations and recruited the patients with Leena Pitkänen. The data analysis was

performed by the author, with assistance by Karhunen-Enckell, Anna Anttalainen and Pekka Kumpulainen. The author led the manuscript writing process and submission in the role of the corresponding author.

1 INTRODUCTION

Cancer accounts for approximately every sixth death in the world [1]. It has remained one of the leading causes of mortality in the world despite decades of research and massive monetary investment. Although the prognosis of many cancer types has improved and continues to improve, the most recent global statistics indicate that the role of cancer as a cause of death will only increase in the future as the global life expectancy is also rising [2]. This ensures that the development of new treatment and diagnostic methods for cancer will remain one of the most relevant topics in the field of medical technology.

Surgery is an important treatment method for many solid cancers, and for some cancer types, such as early-stage breast cancer, it is nearly always part of the primary treatment [3,4]. The goal of surgical treatment is to remove the cancerous mass as comprehensively as possible without compromising vital structures. Furthermore, as the survival rate of cancer patients is constantly improving and patients often live for decades after the surgical operation, an ever-increasing emphasis is also being placed on the functional and aesthetic outcomes of surgery, in consideration of the quality of life and overall health of the patient [5–7]. In the case of breast cancer, improved cosmetic outcomes are associated with smaller resection volumes [8]. Thus, the removal of benign tissue surrounding the cancer should be minimized. However, due to the widespread and irregular nature of cancer, it is often difficult to assess accurately where the cancer ends and benign tissue begins, especially during a surgical procedure. This can lead to an incomplete removal of the cancer – i.e., a positive surgical margin – which in turn can lead to recurrence and a reoperation [9]. The average rate of reoperations for breast cancer is approximately 20%, but it can vary significantly from less than 10% to more than 50%, depending on the region and cancer type [10,11]. The failure to remove the cancer completely in one operation is an important issue, since the need to go through additional surgery can be devastating for the cancer patient in terms of both physical and mental well-being [7,12]. In addition, every reoperation adds to the total healthcare costs [13].

Due to the incentives of overall health and economics, tools and methods to improve intraoperative surgical margin assessment have been introduced in recent

years. However, these methods have not yet fully spread to wider clinical use due to them being inaccurate, slow, inconvenient to use or too expensive. For a method to be widely accepted as an aid in margin detection, it needs to fit the workflow of a surgeon or a pathologist, provide fast and accurate results, and be cost-effective for the hospital.

This thesis presents the concept and realisation of a system in which the negative aspects of the current auxiliary methods are overcome or mitigated by means of tissue identification from surgical smoke by differential mobility spectrometry (DMS). DMS is a measurement technology that utilises a high-frequency asymmetric electric field to separate ionised molecules to provide a spectrum of the molecular content of a gaseous sample. DMS has previously been used in medical research, for example, to identify diseases from human breath samples or bodily fluids [14,15], but this thesis presents the first studies where the technology was used for tissue identification in a surgical application.

1.1 Objective

The objective of this thesis is to introduce a concept for a new *in vitro* medical device that could be used for tissue identification in a clinical environment. More specifically, the aim of the research that is presented in the thesis is to determine whether clinically relevant tissue types can be accurately discriminated from each other based on surgical smoke by using a DMS-based system in a surgical setting. In the thesis, the technological progression of the system is explained, starting from its use in animal tissue studies in a laboratory environment and concluding with its realisation as a medical research device in an actual surgical setting. The ultimate objective is ambitious, and to accomplish it, the feasibility and the key limitations of the system regarding sample acquisition and duration need to be studied and resolved. This is accomplished through five scientific publications, each of which aims to resolve a research question related to the measurement system. The research questions are as follows:

I. Is DMS analysis of surgical smoke a feasible method for *ex vivo* tissue identification in a porcine model?

II. Can DMS-based *ex vivo* tissue identification be utilised with clinically relevant human breast tissue?

III. Is DMS-based *ex vivo* tissue identification feasible as an imaging method for pathological applications?

IV. Can DMS-based tissue identification be performed rapidly enough to enable intraoperative use?

V. Is DMS-based tissue identification feasible in a surgical setting?

The publications and their findings serve a specific purpose to build up the researched system to make it suitable for its intended application of surgical margin assessment and to point out the shortcomings and challenges that the method exhibits. Firstly, in Publication I, the focus is put on the overall feasibility of tissue identification from surgical smoke by DMS. Publication II further expands on the study of feasibility to prove the clinical relevance of the system by identifying benign and malignant human breast tissue. The purpose of Publication III is to explore whether the DMS-based tissue identification could also be incorporated in pathology and thus expand the application area of the methodology. The focus of Publication IV is in the speed of analysis and the requirements for the system that are needed to accommodate real-time use. Lastly, Publication V concentrates on the performance, usability aspects and limitations of the system in a clinical environment and *in vivo* use.

1.2 Contribution

The main scientific contribution of this thesis is the translational research regarding the feasibility of tissue identification with a novel application that combines sensor technology, data science and clinical medicine. The research culminates with the realisation of a prototype system that allows for tissue identification based on DMS analysis of surgical smoke. The properties of the device are introduced in each phase of the development trajectory. The benefits and limitations of the system are discussed and compared to competing methods. Data analysis methods to enable the deduction of tissue types through raw DMS data are also discussed and evaluated. Although the system presented in this thesis is not yet applicable to widespread clinical practice, the technology has the potential to provide a significant positive societal impact by improving the treatment of solid cancers in the future.

1.3 Structure

In this thesis, the information of five scientific journal publications is expanded upon and presented in the following chapters: Literature review and background, Theoretical background, Materials and methods, Results and discussion, and Conclusions.

2 LITERATURE REVIEW AND BACKGROUND

In this chapter, the basis of the medical problem of assessing the surgical margin in cancer is introduced alongside publications that illustrate the importance of resolving the related issues. This chapter also includes a thorough review of the current methods that aim to help in the margin assessment and explains their benefits and disadvantages. The development of a device for surgical use that would overcome the issues of the current methods requires a multidisciplinary approach. To understand the measurement quantities and their connection to the tissue type, some background on medical biology is needed. Furthermore, the measurement principle itself requires an understanding of the fundamentals of differential mobility spectrometry. In addition, the means of translating the measured DMS data back to a tissue type involve some of the core methods of machine learning. These topics are also introduced in this chapter.

2.1 Biological foundation

Mutations within our genome are rare and occur naturally at an estimated rate of three nucleotides (out of three billion) per cell division [16]. Mutations can also occur due to external factors, such as ultraviolet radiation from sunlight or tobacco smoke, that cause damage to the deoxyribonucleic acid (DNA) strands in our cells. Despite the damage, multiple repair mechanisms in the replication and transcription of DNA are able to effectively limit the number changes in the genome to an acceptably low level [17]. In addition, some mutations that pass through the repair mechanism alter the cell in a way that causes it to not survive long enough to cause any additional issues. Still, during our lifetime and several cell generations, the number of random mutations is constantly accumulating, and the older we get, the more likely they are to cause issues [18]. Most often, the mutations have little to no effect on the health or function of an organism, but when the mutations happen in certain parts of the genome, the so-called proto-oncogenes or tumour suppressor genes, the result of the mutation can cause a cell to gain an ability that promotes abnormal cell behaviour or the loss of a function that suppresses such behaviour. These changes in the cell functionality are the root cause of cancer [19].

Cancer describes a disease in which the accumulation of mutations in the DNA causes the cell survival and proliferation of tissue to become uncontrollable [19]. In cancer, the normal apoptosis, i.e. controlled cell death, can be inhibited by the downregulation of, for example, the p53 gene, which causes the cells to survive and multiply for longer than they normally would [20]. However, the excessive survival and proliferation alone do not constitute cancer. If the mutated cell and its progeny lack the ability to invade the surrounding tissues and do not spread through the blood circulation or the lymphatic system, the abnormal tissue growth is categorized as a neoplasm or a benign tumour. Benign tumours are extremely common and can manifest as, for example, moles or fibroids. In this case, the prognosis is often unaffected, and removal of the tumour is not necessarily even needed. Nevertheless, benign tumours have the potential to progress to a malignant form and thus function as a precursor for cancer. The phases of tumour progression are illustrated in Figure 1.

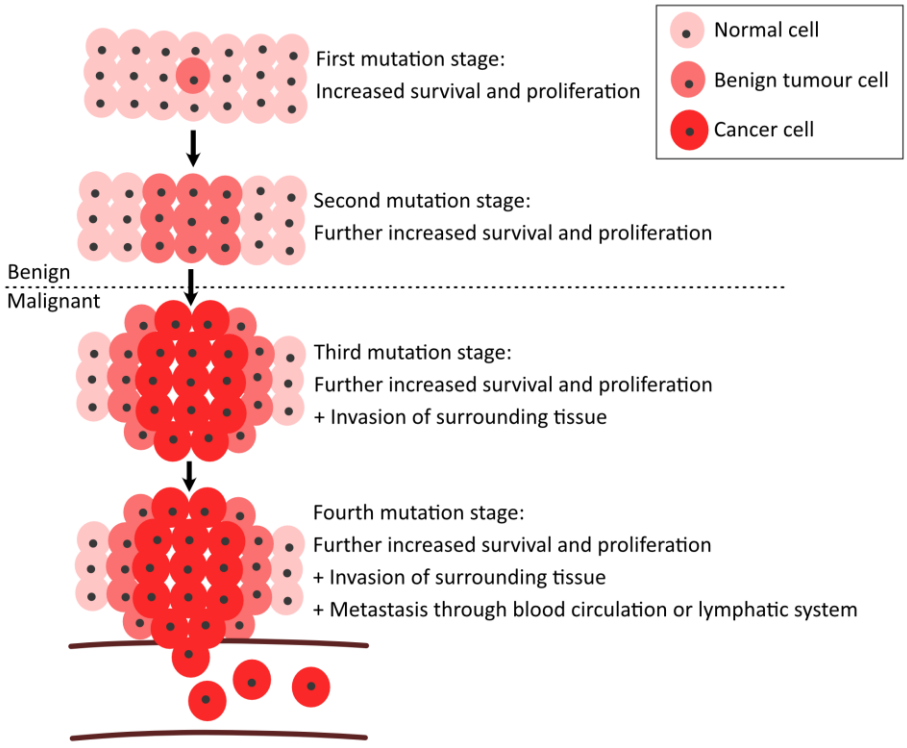


Figure 1. *A simplified example of tumour progression.*

The schematic representation of tumour progression in Figure 1 is simplified, and, in reality, there can be several substages of mutations and the division between a benign and malignant tumour is not always straightforward. A relevant example of this is *ductal carcinoma in situ* (DCIS), a type of neoplasm that is considered malignant, even though the growth is localised to the mammary ducts. DCIS is a disease that has seen a significant increase in incidence over the last three decades. In the 1980s, the disease was a mere curiosity in the statistics, but currently approximately 20%–35% of all breast cancers are diagnosed as DCIS [21]. The radical increase in the incidence is due to cancer screening becoming more common. In terms of its biological properties, DCIS is heterogeneous and difficult to categorize appropriately. Furthermore, while the disease itself has a good prognosis, it already contains almost all of the genetic instabilities that enable its progression. The aspects surrounding its evolution to an invasive carcinoma, such as the proportion of cases and progression timelines and mechanisms, are still largely unknown [22]. In addition, because DCIS grows inside the ducts, it is hard to assess the tumour during surgery, since it is not macroscopically visible and is rarely palpable [23]. The inherent complexity, elusiveness in a physical examination and unpredictability of progression are the reasons why the early and accurate diagnosis of DCIS is crucial and why the disease encapsulates the need for improved methods and research in cancer diagnosis and surgical treatment.

However, DCIS is only one example of the diversity of cancer types. Cancer can originate from practically any organ and cell type, and the differences in their origin and distinctive traits affect their incidence, prognosis and treatment options [1,2]. For example, in some cases of leukaemia, i.e. cancers that originate from blood-forming tissue, the surgical excision of a distinct tumour mass is not possible, and the treatment is primarily medication-based. On the other hand, for most solid tumours, surgical treatment is an effective treatment method that cannot be replaced with medication (i.e. chemotherapy) alone since, the blood circulation through the tissue is impaired. Despite the different approaches and vast variety of cancers, malignant tumours share similar traits in terms of their cellular structure and metabolism.

2.1.1 Biomarkers

Each cell type has a characteristic molecular signature, which defines its structure and function. Almost the entirety of the dry mass of any cell comes from four types of organic molecules: nucleic acids, carbohydrates, proteins and lipids [24].

However, their relative abundance in normal tissue cells varies depending on the degree of differentiation of the cell and the tissue type. For example, in adipose tissue, the mass of a fully differentiated cell consists primarily of lipids, whereas in skeletal muscle, water and proteins are the dominating molecules [25]. Similar, albeit less profound, differences exist between malignant and benign cells, even when they originate from the same tissue. This is partly explained by the fact that, in cancer cells, the degree of cell differentiation is not as straightforward as in benign cells and they can, in fact, even de-differentiate to resemble a non-differentiated cell [26]. In general, the more differentiated a cancer cell is – that is to say the more it resembles a surrounding normal healthy cell – the better the prognosis [26]. In cancer cells, the characteristic biomolecules that differ in their content or relative abundance from benign cells are what enable the identification of the disease and can thus be considered cancer *biomarkers* [27]. Biomarkers in cancer can be identified throughout their progression from normal healthy tissue to malignancy. Based on their type, biomarkers are sometimes complemented with the terms genetic, molecular and metabolic. For example, known tumour-suppressing genes or proto-oncogenes can be considered genetic biomarkers. One example of a genetic biomarker is the BRCA1 gene, where an abnormality leads to the development of breast cancer by the age of 80 in approximately 72% of cases [28]. The gene maintains the normal repairing process of DNA by encoding the breast cancer type 1 protein, which in itself can be considered a molecular biomarker. The depleted expression of the protein leads to abnormal cellular responses to damaged DNA, and among numerous other effects, this increases the synthesis of certain lipids (specific examples can be found in Table 1, p. 30) [29]. The overabundance of lipids, then, can be considered a metabolic biomarker. Thus, the same mutation can lead to multiple different ways of identifying its existence by utilising different biomarkers (Figure 2).

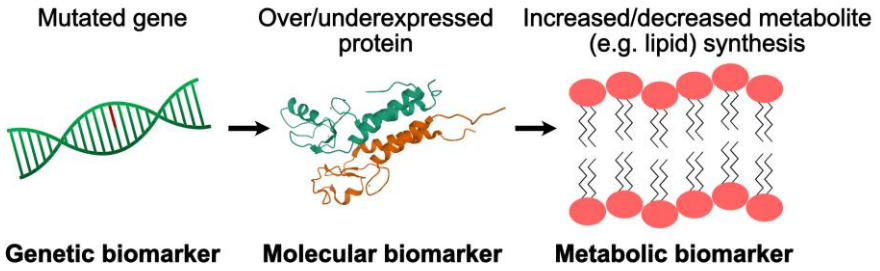


Figure 2. *Types of biomarkers.*

While the term biomarker can be used to describe any biomolecule that can be objectively measured and used to indicate normal biological or pathological processes or the risk to develop a disease, the majority of hypothesized and studied biomarkers are not approved for clinical use [30]. For example, the number of proteins that are approved by the US Food and Drug Administration as biomarkers for cancer is limited, even though studies have listed thousands of proteins that have exhibited a clear link to hundreds of different cancers [31,32]. The rigorous process of providing statistically significant evidence of the connection between a certain biomolecule and its relative abundance in cancer is time-consuming and taxing on resources. Furthermore, specific threshold limits that separate cancer from benign tissue with adequate sensitivity and specificity are not easily determined for a single biomolecule. Thus, modern biomarker research has focused on finding a connection to cancer based on a combination of biomolecules so that the diagnosis relies on a more comprehensive set of relevant biomolecules rather than a singular threshold value [33]. The scientific fields in biology that have been described with the collective term omics have been at the forefront of discovering new potential biomarkers for cancer. Omics comprises the disciplines that study the possible sources of biomarkers, such as proteins (proteomics), genes (genomics) or metabolic products (metabolomics). Omics is also relevant in this thesis, since the proposed measurement system is designed to measure the vaporised cell contents that contain these biomolecules and their degradation products. By basing the tissue identification on the normal metabolites of the cell that are present in every cell type, the system cannot measure one definitive marker that would automatically determine the malignancy or type of the tissue. Rather, the differentiation comes from the relative abundances of these substances and the resulting differences in their molecular fingerprint. The lack of specific markers creates an overlap between the measured outputs of analysed tissues, which can limit the differentiation capability. On the other hand, the general nature of the utilized biomarkers does not restrict the method to a single use case, since almost all cancers exhibit a specific metabolic abnormality.

2.1.2 Warburg effect

The genetic and molecular biomarkers that are connected to different types of cancers are rarely the same, and it is obvious that, for example, a biomarker called “prostate specific antigen” is clinically relevant in the screening and indication of prostate cancer but has no value in the treatment or diagnosis of, for instance, brain cancer. However, a certain metabolic pathway that was discovered almost a century ago has

been proven to be very generalizable to all kinds of malignant tumours. The pathway in question is called the Warburg effect, named after its discoverer [34].

The Warburg effect describes a metabolic abnormality in which a cell favours a more inefficient pathway to produce energy, even when normal metabolism is not restricted by the lack of oxygen in the cell. More specifically, the cells that exhibit the Warburg effect produce most of their energy units, adenosine 5′-triphosphates, through aerobic glycolysis instead of normal mitochondrial oxidative phosphorylation, even though the resulting energy yield is approximately nine times lower [35]. Due to the Warburg effect, cancer cells can have a significantly increased glucose uptake and subsequent lactate production. Figure 3 shows a simplified depiction of the phenomenon.

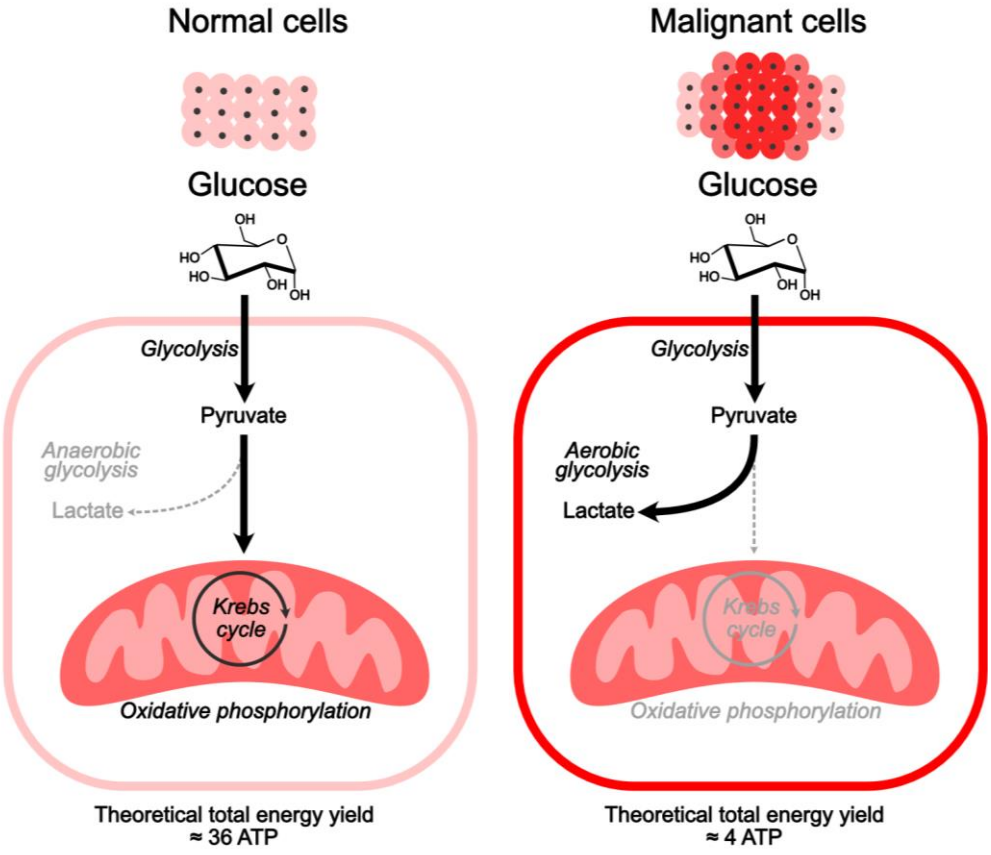


Figure 3. *The Warburg effect.*

Unlike Warburg himself hypothesised, this abnormal behaviour does not transpire due to cell respiratory damage, but rather since it seems to be beneficial for the proliferation and survival of the cell [36,37]. There are still unanswered questions regarding the reasons for and benefits of the Warburg effect for cancer cells, but it has been theorised that the high rate of glycolysis and the resulting excess of glycolytic intermediate products prime the cell to activate biosynthetic pathways to accommodate for the rapid proliferation, when required [37,38]. Thus, cancer cells can prioritise the accumulation of biomass over efficient energy production [35]. It has also been suggested that the Warburg effect can explain all or most of the hallmarks of cancer, such as the invasiveness, inhibited apoptosis and the uncontrolled proliferation [39]. That being said, some studies contradict this and indicate that the Warburg effect is not universal to all cancer cell lines and the current understanding is that the diversity of metabolic activity is what actually benefits the rapid proliferation of cancer cells [37,40]. Although the Warburg effect and its extent in cancer cells is not yet fully understood, it is likely a significant contributing factor for the irregular molecular composition of cancer cells. The over-abundance of the intermediate products of biomolecules, such as amino acids, nucleic acids and lipids, arguably suggests the increased biosynthesis of some of these molecules [41]. In fact, several lipids and lipoproteins have been shown to be upregulated (or downregulated) in various types of cancers [42]. This type of change in the relative abundance of the biomolecules can make cancer tissue identifiable at a molecular level. As an example, studies have shown that the content profiles of the cell membranes of different cancerous cell lines are very similar for many specific phospholipids [43]. Phospholipids are biomolecules that are important in forming the cell membranes and make up approximately 3% of the total mass of a mammalian cell [44]. They consist of two fatty acid chains that form a hydrophobic tail and of a hydrophilic head comprising a glycerol group, phosphate group and a small polar group, such as serine or choline. Due to the possible variation in the length of the fatty acid hydrocarbon chains, in the number of double bonds and in the polar head group, cell membranes can consist of as many as 500–1000 different phospholipid species [45]. Table 1 shows the relative fold change of some of these phospholipids, along with certain sphingolipids, between normal and cancerous breast tissue. The abundance of all measured lipids, except for sphingomyelins, was increased in tumour tissue in the presented groups in the referenced study, where the comparison was made between normal ductal breast tissue and breast cancers [46].

Table 1. *Comparison of lipid groups in cancerous and normal breast tissue* [46]. *N refers to the number of different species from a certain group.*

Lipid class	Lipid group	N	Fold change for tumour vs normal tissue
Phospholipids	Phosphatidylcholines	23	2.85–34.30
	Phosphatidylethanolamines	18	2.87–75.06
	Phosphatidylinositols	2	3.32–10.34
Sphingolipids	Sphingomyelins	8	0.50–6.36
	Ceramides	2	2.31–5.62
	Glucosylceramides	4	2.40–11.63

However, as stated before, the difference in a few selected biomarkers might not be the best basis for tissue identification due to the difficulty in determining fixed threshold levels for benign and malignant tissue. For example, the menstruation cycle can create variability in the lipid metabolism of normal benign breast cells [47]. In breast cancer cells, the grade (i.e. degree of differentiation) of the tumour, hormone receptor status and some medications can affect metabolism and the lipid composition [46,48]. In addition, the total lipid content of cells has been shown to vary significantly and increase the closer the cells are to the edge of the tumour [49]. Similar behaviour has even been reported in peritumoral benign tissue [49,50]. This type of variation in specific biomolecules supports the idea that more effective tissue identification could be achieved by analysing the full molecular fingerprint or other more general properties of the sample rather than concentrating on specific biomarkers.

2.2 Surgical margin assessment

The metabolic and cellular biomarkers between cancerous and benign tissue can enable the differentiation of the tissues, especially when there is a possibility for microscopic examination. However, at a macroscopic level, the differences are not always clear and a visual examination and palpation alone do not necessarily suffice.

Still, in the surgical treatment of cancer, the success of removing the entirety of the tumour mass is often entirely dependent on the experience and visual assessment of the operating surgeon and guidance from imaging performed prior to surgery. The aim of the surgeon is generally to excise the malignant tissue with minimal damage to the patient and the surrounding benign tissue. Therefore, the operation becomes an optimisation problem, where the measure of success is defined through the benign tissue that surrounds the excised tumour, i.e. the *surgical margin* or *resection margin*. To define the adequacy and categorisation of surgical margins, there are several dedicated regional and international guidelines for different types of cancers and their respective tumour grades [51–53], but, in general, the outcome of surgical margin assessment can be considered binary. In a setting where the tumour has not infiltrated vital structures and can be completely removed, the margin is defined as *positive*, if there are cancer cells near or within the border of the excised tumour, and as *negative* or *clear*, if the tumour is surrounded by a certain width of benign tissue in all directions [54,55]. If a negative margin cannot be achieved, the risk of cancer recurrence through residual cancer cells is elevated and additional treatment is required [54]. The additional treatment often involves a reoperation, which adds to the total healthcare costs and burdens the patient, which is why negative margins are a central aim in surgical treatment.

The importance of negative margins and their accuracy is accentuated in the case of breast cancer. The reason for this is that breast cancers can often be removed in their entirety unlike, most brain cancers, for example, that infiltrate the benign tissue more unpredictably, resulting in a more sporadic shape and a standard of care that emphasises minimizing the functional deficit of the patient over clear surgical margins [56,57]. In addition, breast cancer is the most common cancer type in women, and the primary treatment of breast cancer almost always consists of surgical treatment, unless the cancer has progressed to a terminal stage [3,4]. In numbers, this translates to over two million cases per year, most of which are treated with surgery [1,58]. Furthermore, the prognosis for early-stage breast cancer is extremely good, with over 90% of the patients alive after five years from diagnosis [59]. This means that most of the patients have a long life ahead of them after cancer surgery, which increases the importance of the cosmetic and functional outcomes of the resection due to their direct impact on the quality of life and overall health of the patient [5,6]. Studies have also shown that smaller resection volumes result in better cosmetic outcomes and that the full removal of the breast, i.e. mastectomy, does not improve the prognosis or decrease the morbidity compared to breast-conserving treatment

[8,60–62]. Therefore, surgical margin optimisation is particularly relevant for breast cancers.

For invasive breast cancer, the current European and American guidelines regarding surgical margins define a negative margin with the “no ink on tumour” rule [4,9,63]. This means that if the outer surface of the resected tumour has no cancer cells that are visible in a histopathological examination after the administration of a special ink that stains the cells, the margin is deemed negative. In the case of DCIS, a negative margin requires no cancer cells to be present within a 2 mm distance from the edge [64]. However, the width of the realised margin is always decided by the operating surgeon, and since it is hard to assess the margin visually during the surgery, some surgeons prefer to be safe and leave a much larger margin, while others try to spare as much benign tissue as possible [65]. As both approaches have their risks and benefits, the personal preference of a surgeon is justifiable, but the guidelines aim to offer a balanced decision point between minimizing the probability of recurrence and the unnecessary loss of benign tissue. A generalized depiction of the optimisation problem related to surgical margins is presented in Figure 4.

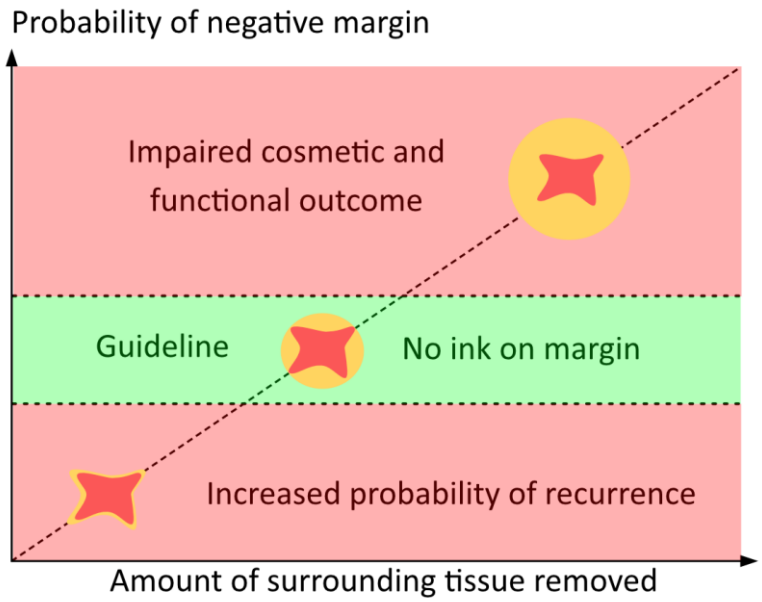


Figure 4. *The balancing act of surgical margins. The dark red shape represents the tumour, and the yellow shape represents the surrounding benign tissue being removed. The background colours represent favourable (green) and unfavourable (red) outcomes.*

Differences in the experience, technical skill and personal preferences of surgeons likely explain the varying rate of reoperations among institutions and regions due to positive margins in breast cancer surgery. In Europe and the USA, the goal for the rate of reoperations has been stated as being 10% and less than 20%, respectively [66,67]. However, based on multiple studies, the reported rates vary from less than 10% to more than 50% [68–71]. The high reoperation rate due to positive margins in some institutions is a significant issue, since reoperations are a detriment to the patients and healthcare system in many ways. The need for a reoperation can worsen the prognosis due to delays in starting adjuvant chemotherapy [12]. In addition, the reoperation can cause severe stress and worry for the patient and increase the cosmetic damage [7]. Reoperations also add to the total healthcare costs of the treatment, and studies have shown that each reoperation can add a cost of several thousands of euros [13]. In a Canadian study, it was estimated that a reduction in the reoperation rate from 23% to 10% would bring a saving of USD 1055 per patient and an annual savings of USD 1.9 million in the region of British Columbia alone [72]. If similar reduction could be achieved globally, the estimated maximum annual savings would be over USD 3 billion.

The research community and medical device developers have acknowledged the importance and challenges of acquiring clear surgical margins, and a number of devices to aid in margin assessment have thus emerged in recent years. However, the currently existing methods have minor or major issues that have prevented them from becoming widespread among medical experts or researchers. An American study reported that, despite the clear demand for lowering the rate of reoperations, any type of intraoperative pathologic margin assessment is conducted by less than 22% of surgeons, and only 3% utilise a margin assessment device [11].

2.2.1 Overview of margin assessment methods

The poor utilisation rate of intraoperative margin assessment tools is universal. Even the methods that have been accepted as standard of care in some facilities the US have achieved limited national adoption, with a utilisation rate of only a few percent among surgeons [11]. A gross examination by visual inspection and palpation is the most widely used method because of its simplicity. However, the method suffers from poor diagnostic performance, as exemplified by a retrospective study that reported a sensitivity of 49% and specificity of 83% [73]. Besides gross examination, most surgeons rely only on surgical margin assessment performed after the surgery with

the gold standard of a histopathological analysis by a surgical pathologist. The process involves the fixation, sectioning, staining by haematoxylin and eosin, and microscopic visual assessment of the resected tumour. The process is time-consuming, and even though the results are accurate, the final information of possible margin positivity is not often available until several days after the surgery. This means that fast corrective action is no longer possible, and a positive or close finding often results in a reoperation.

To keep the high diagnostic accuracy of histopathological evaluation while still retaining the possibility of intraoperative assessment, a method known as frozen section analysis (FSA) was proposed as early as over a century ago, and it is still in use as standard of care in its birth institution (Mayo Clinic, USA) [74,75]. In FSA, the resected tumour is sectioned into small pieces, which are snap-frozen, stained and delivered for pathological analysis while the patient is under anaesthesia. Based on two separate meta-analyses, FSA has a reported sensitivity of 81%–86% and specificity of 95%–97%, and it has been shown to reduce reoperation rates from 55.3% to 19.3% in a one-year surveillance study [76–78]. However, for an intraoperative method, the analysis is relatively slow, adding a delay of 20–30 minutes to the surgery [79]. In addition, the method is resource-intensive, since it requires laboratory space and trained personnel for the sectioning and analysis. Furthermore, the frozen sectioning can cause damage to the tissue, making traditional post-operative analysis more difficult [80]. Also, a comprehensive microscopic analysis is impossible to conduct while the patient is anesthetised, meaning that only a small fraction of the margin is assessed intraoperatively.

A more comprehensive pathological method for intraoperative margin analysis is imprint cytology. In imprint cytology, the surface on each side of the resected tumour is pressed against glass slides, which are then fixed and stained. The method is based on the idea that only the malignant cells adhere to the slides, while the benign adipose tissue will not leave an imprint [81]. The diagnostic performance of the method is good, with a reported sensitivity of 91% and specificity of 95% in a meta-analysis [76]. However, the method requires an expert cytologist for the analysis, and the delays in the operation are more than ten minutes. In addition, atypical cells that are not malignant but adhere to the cytology slides impede the diagnosis, further emphasising the required experience and expertise of the cytologist [82,83]. While the method has its issues, imprint cytology has been reported to lower reoperation rates from 26% to 4% in a systematic review [79]. It has also been shown that

combining imprint cytology with FSA mitigates the respective downsides of both methods and reduces the positive margin rate even further [84].

Still, regardless of the means of implementation, pathological methods for intraoperative margin assessment are always relatively slow and resource-intensive. Therefore, other methods have been suggested and implemented in clinical practice. For example, various intraoperative imaging methods have been adapted to the surgical workflow at some capacity. Specimen radiography is one of the most common methods and is used in, for example, breast cancer operations in Finland. In specimen radiography, the resected tumour is x-ray imaged, after which the image is assessed by a radiologist. Depending on whether the x-ray imaging is done in the operating room by means of digital specimen radiography or more traditionally in a separate radiology laboratory, the duration of the analysis is approximately one minute or more than twenty minutes, respectively [85]. Specimen radiography has been shown to reduce the rate of positive margins, but, overall, it suffers from poor diagnostic performance, with a reported sensitivity of 53% and specificity of 84% in meta-analyses [76,86,87].

Another imaging method for intraoperative margin assessment is high-frequency ultrasound imaging. In this method, the resected tumour is imaged from all sides by an ultrasound probe in the operating room environment, and additional resections are made if the sonographic margin is positive [88]. Compared to pathology, the analysis is faster, but the assessment is still traditionally carried out *ex vivo* (outside the patient) [89]. This means that the analysis requires an additional tool and that the orientation of the tumour in respect to the resection area must be accurately indicated with guidewires or markings. In a meta-analysis, the sensitivity of the method was reported to be 59% and specificity 81% [76]. Besides specimen radiography and intraoperative ultrasound, some less common imaging methods, such as magnetic resonance imaging and nuclear medicine have, also been applied to margin assessment, but the high resource needs and long duration of the analysis have limited their use, even in the research field [90,91].

Several optical methods have been introduced to and adopted in clinical practice. The optical methods can be used to cover the whole specimen at the expense of spatial resolution, or they can focus on selected areas, which, in turn, decreases the comprehensiveness of the margin analysis. Optical imaging can provide information on the resection surface, but the penetration depth of photons at optical wavelengths is shallow, meaning that deeper structures cannot be imaged [92]. An example of whole specimen optical imaging is photoacoustic tomography, in which nanosecond

laser pulses are emitted to a resected tumour, which causes a differential expansion of molecules, such as lipids, within the tissue and creates ultrasonic signatures that can be imaged [92]. The diagnostic performance of photoacoustic tomography has been promising, with reported results of over 90% for both sensitivity and specificity at their best, but the results have high variance depending on the interpreter [92,93]. Similar results have also been achieved with other whole specimen optical imaging methods, such as fluorescence imaging [94]. As for optical imaging methods designed for the assessment of smaller areas of the sample, optical coherence tomography and Raman spectroscopy are examples that have been studied in the context of margin assessment. In optical coherence tomography, near-infrared light is used to create high-resolution (2-10 µm) images of the tumour surface at a depth of approximately 2 mm, based on reflected light from the tissue [95,96]. Raman spectroscopy, on the other hand, relies on detecting differing molecular vibrations from the tissue after excitation with monochromatic light [92]. Both methods have resulted in good diagnostic performance in tissue differentiation, with reported sensitivity and specificity of around 90% [97,98]. However, much like other optical methods, their widespread use in intraoperative margin assessment in clinical care has been limited due the downsides of the methods. The most pressing issues with optical methods in intraoperative margin detection are the need for an experienced interpretation of the imaging results, a limited imaging depth and, in most cases, the requirement of *ex vivo* analysis and the associated delays.

Among the device-based margin assessment methods, radiofrequency spectroscopy has perhaps been the most successful in entering clinical practice. The MarginProbe™ device based on this method has been approved for clinical use in the USA and the EU, among other regions. The method relies on detecting differences in the electrical properties of resected tissue surfaces with a hand-held probe. The structural and molecular differences in benign and malignant tissues affect the electromagnetic scattering, absorbance and reflectance of the tissue and can be detected with radiofrequency spectroscopy [99]. The expected duration of the margin examination is within the range of a few minutes but depends on the thoroughness of the analysis and the surface area of the tumour. The downsides of the method are its reliance on *ex vivo* analysis with a separate tool and the relatively poor reported diagnostic performance with an average sensitivity of 70% and specificity of 59% [76]. Even with the suboptimal diagnostic performance, radiofrequency spectroscopy has been reported to lower reoperation rates in randomized trials [99,100]. However, at least in one retrospective study, a statistically significant improvement could not be inferred [11].

2.2.2 Molecular margin assessment methods

Among the surgical margin assessment methods, the subgroup that utilises molecular detection bears the most relevance to the approach proposed in this thesis. The operation principle of these methods is rooted in the cellular and molecular differences between cancer and benign tissue and in the detection of these substances with high-precision instruments, or, more specifically, mass spectrometers (MS). Mass spectrometry (MS) relies on the ionisation of gas-phase analytes and the separation and detection of the analyte ions based on their mass-to-charge ratio. MS is a versatile and accurate analysis method that can produce quantifiable information on the molecules that are present in the sample [101].

MS has been utilised as a measurement technology in intraoperative tissue identification with various sampling methods. One of the newer methods, introduced in 2017, utilises water (or other biocompatible solvent) droplets (4–20 µl) to extract biomolecules from the tissue surface by a specialised sampling probe [102]. After contact with the tissue, the droplet is extracted into an MS instrument by its vacuum for molecular analysis. The system, named MasSpec pen, is fast, with one measurement taking only approximately ten seconds. The performance in the *ex vivo* identification of various cancer types from benign tissue has also been very high, with results for sensitivity and specificity being generally well over 90% [102,103]. The technology has also been piloted in intraoperative use for *in vivo* analysis, but the diagnostic performance was not reported in the study [104]. As its advantage, the method does not damage the analysed tissue at all. Still, the need for a separate probe reduces the compatibility with normal surgical workflow. However, in robot-assisted surgery, as an additional drop-in probe, the system could bring added value without limiting the workflow, and it has been piloted in a porcine model as a part of automated surgery with excellent results as regards tissue identification [105].

MS-based intraoperative tissue identification has also been utilised in methods that rely on the ablation of tissue with a laser probe. The introduced methods differ slightly in terms of their operation principle, but, for example, carbon dioxide lasers and solid-state lasers with nano- and picosecond pulsing have been used [106–108]. Commonly, the laser wavelength in the methods has been in the mid-infrared (3–6 µm) range to optimise tissue absorbance. The tissue damage due to the laser ablation is somewhat dependent on the specifications, but, in general, the methods are minimally destructive, with the depth and diameter of the sampling area being less than one millimetre [108,109]. The methods are relatively fast, with irradiation times of approximately 10 seconds for the areas of interest. The tissue identification results

with the laser-based methods have been over 95% for specificity and sensitivity in the benign versus malignant tissue setting, and the methods have even demonstrated their capability in cancer grade and subtype identification [110,111]. However, as with the MasSpec pen, the need for an additional probe in the laser ablation methods is a disadvantage in terms of usability.

One MS-based method that can be incorporated into the normal surgical workflow is rapid evaporative ionization mass spectrometry (REIMS). REIMS has been in research use for more than a decade, and while it was initially conceived for the purpose of *ex vivo* and *in vivo* tissue identification in the field of medicine, the application area of the method has expanded even to agriculture and the food industry [112–114]. The operation principle of REIMS is similar to the system introduced in this thesis, as it uses an energy instrument to produce a gaseous sample material that is subsequently delivered to a gas analyser. Most commonly, the energy instrument used with the method is standard electrosurgery. As the method has been researched for several years, tissue identification results for various cancers have been reported, including liver, skin, cervical, rectal and breast cancers [115–119]. In general, the results have been excellent, with sensitivities and specificities nearing 100%, especially in *ex vivo* studies. *In vivo* diagnostic results are rarely reported and do not reach the level of the *ex vivo* results [116,118]. One likely reason is the difficulty in intraoperative annotation and the resulting weak labels for samples. REIMS analysis is fast, with a duration of less than three seconds from sample introduction to result presentation [119,120]. Compared to the other MS-based methods, REIMS is destructive by nature, since it is directly connected to the surgical incision, but the method itself does not cause any additional tissue damage. Figure 5 shows some of the different MS-based sampling methods along with their characteristics.

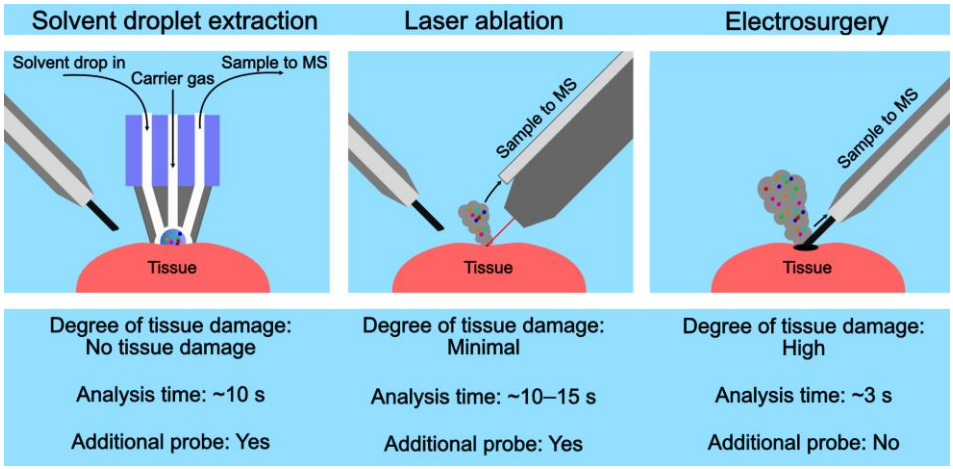


Figure 5. *Sampling characteristics of intraoperative tissue identification methods based on mass spectrometry (MS).*

The MS-based methods are an efficient way to identify tissue malignancies due to their extreme precision to detect trace amounts of biomolecules that are integral to the formation of cancer. In all of the methods, the observed key difference between cancerous and benign tissue in the published studies has been observed in the relative abundance of different lipids, such as phospholipids and triglycerides [102,106–108,119]. Thus, the methods rely on the inherent and generalisable metabolic traits of cancer, such as the Warburg effect. One REIMS study has even shown that the identification of breast cancer is possible by training a classification model with the MS features of skin cancer [121]. The general nature of the identification and simultaneous utilisation of signals from several biomolecules increase the number of potential applications and are one reason for the high diagnostic performance.

Although all of the methods based on MS are accurate and their use is either minimally intrusive or fits perfectly in the surgical workflow, MS is a high-end analysis device that requires a high vacuum to function. This serves as an inhibiting factor for widespread use of the methods. Mass spectrometers have high initial and upkeep costs (up to USD 1 million and 10% of purchase price annually), and their maintenance requires an educated expert [122]. In addition, MS devices require frequent, even day-to-day, calibration and tuning, and specific reference libraries [123]. High-end MS instruments are also large (> 1 m^3, > 150 kg), which makes their relocation and overall implementation in operating rooms more difficult. However, as research instruments, the potential of MS-based methods in tissue identification is unmatched.

Overall, the field of intraoperative tissue identification has seen various innovative solutions that have produced promising results in scientific studies. Still, the clinical adoption of these methods has been limited, and a real breakthrough method that could resolve the issues of the current technologies is yet to emerge. The DMS-based solution presented in this thesis aims to fulfil this need. For comparison, Table 2 presents the strengths and weaknesses of selected methods for intraoperative margin detection. Emphasis must be placed on the fact that the values in Table 2 represent the subjective assessment of the author, based on the research literature reviewed at the time of writing this thesis, and can be disputable in some parts.

Table 2. *Comparison of selected intraoperative margin assessment methods.*

Method	Time	Area	Compatibility with workflow	Performance (3 x in Total)	Resource needs	Size	Clinical adoption	Total
Visual examination and palpation	3	3	3	1 (3)	3	3	3	21
Frozen section analysis	0	0	0	3 (9)	0	0	2	11
Imprint cytology	0	1	0	3 (9)	0	0	2	12
Specimen radiography	0	0	0	1 (3)	0	0	3	6
Ultrasound imaging	1	2	1	1 (3)	1	2	2	12
Optical methods	1	2	1	3 (9)	1	2	2	17
Radiofrequency spectroscopy	1	2	1	1 (3)	3	3	2	15
MasSpec pen	2	3	2	3 (9)	2	1	0	19
Laser ablation MS methods	2	3	2	3 (9)	2	1	0	19
REIMS	3	3	3	3 (9)	2	1	1	22
DMS-based method (current)	2	3	3	1 (3)	3	2	0	16
DMS-based method (aim)	3	3	3	3 (9)	3	3	3	27

Time: Result from examined area; 3 = less than 5 seconds, 2 = less than 15 seconds, 1 = less than 5 minutes, 0 = more than 5 minutes. **Area:** Percentage of margin assessed in intended use; 3 = 100%, 2 = more than 50%, 1 = more than 10 %, 0 = less than 10%. **Compatibility with workflow:** 3 = requires no additional procedures, 2 = requires an additional tool (*in vivo*), 1 = requires an additional tool (*ex vivo* analysis), 0 = requires transfer of specimen outside operating room. **Performance (valued as three times as important as other metrics):** 3 = diagnostic accuracy > 90%, 2 = diagnostic accuracy > 80%, 1 = diagnostic accuracy > 60%, 0 = diagnostic accuracy < 60%. **Resource needs:** 3 = no additional personnel needed, low-maintenance; 2 = no additional personnel needed, high-maintenance; 1 = result interpretation needs specific training or additional personnel; 0 = need for separate laboratory analysis. **Size:** 3 = standard operating room instrument or smaller (e.g., surgical evacuator or energy surgery unit), 2 = larger than two standard operating room instruments, 1 = larger than three standard operating room instruments, 0 = need for separate facility. **Clinical adoption:** 3 = in standard clinical use worldwide, 2 = in clinical use in some institutions, 1 = in commercial research use, 0 = only used by one or few research groups.

2.3 Surgical smoke

As described earlier, the assessment of surgical margins is integral to the treatment of solid tumours, and new technological solutions for the problem could be beneficial. However, to increase the likelihood of clinical adoption, the assessment method should be compatible with the surgical workflow. This means that an ideal auxiliary system for margin assessment should utilise information that is produced naturally during every surgery. This can be achieved by means of analysing of surgical smoke. Surgical smoke is created as a by-product in all energy instrument surgeries. Surgical energy instruments are based on heating the target tissue with, for example, electromagnetic radiation (laser), ultrasound-induced mechanical vibration or electricity. The energy modality has been shown to affect the particle size and composition of the produced surgical smoke due to differences in the amount of transferred energy and subsequent temperature elevation [124–126]. The system presented in this thesis has thus far only been used to analyse surgical smoke produced by standard monopolar electrosurgery but will likely be validated for other energy instruments as well, since their role in general and cancer surgery will increase in the future [127].

Standard electrosurgery, also known as diathermy, is routinely used in the excision of solid tumours. In electrosurgery, alternating current with a frequency between 200 kHz and 3.3 MHz evaporates tissue to form an incision or coagulates the tissue to achieve haemostasis [128]. Evaporation occurs, when the operated tissue is rapidly heated by an electrical arc with a continuous sinusoidal waveform and roughly 1 kV peak-to-peak voltage. In standard electrosurgical units, this is usually achieved with the *cut* mode. In the c*oagulation* mode (often shortened to *coag*), the voltage is higher, approximately 3–5 kV peak-to-peak, and it is conducted to the tissue with a low duty cycle. As a result, the temperature increase in the tissue is higher and the thermal spread wider, and, in addition to tissue evaporation, the non-vaporised tissue forms a charred black layer onto its immediate surface, i.e., fulgurates. If the surgical electrode forms direct contact with tissue, the temperature elevation is slower, and with a short exposure, the tissue does not evaporate but rather desiccates, meaning that the cells essentially dry out without extensive molecular decomposition. The result of electrosurgical cuts can also be affected by modifying the speed of incisions and by blending the cutting modes by altering the duty cycle and output voltage [128,129]. Figure 6 depicts different cutting modes and their respective characteristics. The mode of electrosurgery has been shown to affect the composition of surgical smoke in previous REIMS measurements, with the cut mode

producing higher signal peaks for phospholipids and the coagulation mode producing higher peaks for triglycerides [130]. In the research presented in this thesis, the electrosurgery mode was kept consistently in cut mode to minimize sampling-related variations in the feasibility studies, but further research is required to validate the method for different modes.

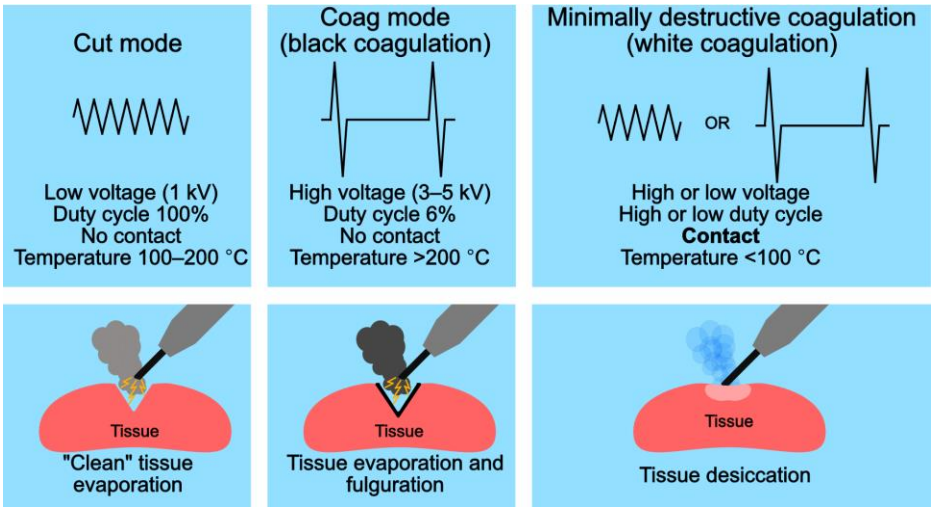


Figure 6. *Standard electrosurgery practices and their effect on tissue.*

The surgical smoke formed in electrosurgery is harmful for the operating room staff through the inhalation of the produced particulate matter [131,132]. There are not many comprehensive studies regarding the general composition of electrosurgical smoke, but it has been stated that 5% of the smoke consists of vaporised cellular content possibly including structural biomolecules, blood, viruses and bacteria, and the remaining 95% is water or steam [133]. Of the vaporised cellular contents of surgical smoke, some studies have concentrated on determining the components that are critical to occupational safety and harmful to inhale [134–136]. In these studies, carcinogenic volatile organic compounds (VOCs), such as benzene, toluene and xylene, have been shown to be abundant in surgical smoke. Due to the safety concerns regarding surgical smoke, the emphasis on removing it from the vicinity of hospital staff has increased, and it has been recommended that smoke evacuation should be implemented as standard practice in all smoke-producing energy instrument operations [137]. This, in turn, would increase the applicability of surgical-smoke-based tissue identification to include a wider range of cancer surgeries.

2.4 Differential mobility spectrometry

To identify tissues based on surgical smoke, measuring of the molecular content of the vaporised tissue is a viable option. MS has been shown to be applicable as a measurement instrument, but its complexity, size and high initial and upkeep costs hinder or prevent its clinical adoption. The complexity of MS devices is mainly a result of the need for a vacuum in which a gaseous substance can be directly ionised and measured based on its mass and charge [101]. In atmospheric pressure, gas-phase analytes can be measured based on their ion mobility that represents the speed with which an ionised sample swarm travels in a carrier gas in an electric field [138]. Ion mobility spectrometry (IMS) considers the shape, size and charge of the sample molecules and thus provides partly complementary information compared to MS. However, the resolving power of IMS technology (in the range of $\sim 10^1$–10^2) is several magnitudes lower than that of MS (depending on the operation principle $\sim 10^4$–10^5) [139]. The resolving power measures the capability of an instrument to distinguish adjacent signal peaks (i.e., ions with a similar collisional cross section in IMS or ions with similar mass in MS) from each other [140]. In practical terms, this means that MS is far more effective in separating distinct signal peaks from complex sample matrices, while in IMS different ions can form clusters with similar mobility properties that will more easily overlap in the produced output data. The signal peaks are also wider in IMS than in MS. This difference is mainly explained by the fact that, due to the vacuum in MS, diffusion does not broaden the trajectory of ion packets, as it does in IMS [139]. The vacuum also explains the difference in the clustering of ion species, since the cluster-forming interactions with the analyte and a carrier gas are absent in MS. To compensate for the clustering of ion species in atmospheric pressure, and to increase the ion separation, a technology was introduced by a Russian research group in the 1980s (and made globally public in the 1990s after the Cold War) [141]. They found out that the mobility of an ion swarm is nonlinearly dependent on the strength of the imposed electric field and that the phenomenon can be effectively utilised in separation by subjecting the ion swarms to a radiofrequency (> 200 kHz) asymmetric electric field. This technology is known as differential mobility spectrometry (DMS), often also referred to as field asymmetric ion mobility spectrometry (FAIMS).

The operating principle of DMS instruments can be divided into three main entities: ionisation, filtration and detection. The ionisation of the gas sample can be achieved in multiple ways and with different ionisation energies which affect the total ion yield and what substances are ionised [142]. Common means for ionisation are

Ultraviolet lamps (low ionisation energy, ~10 eV), radioactive isotopes (medium or high energy depending on the source, keV–MeV range), corona discharge (high ionisation energy, MeV range), and x-ray sources (medium ionisation energy, keV range). Especially with medium and high ionisation energies, the analyte ions are not normally the primary ion species in the ionised gas. Instead, due to the extreme abundance of carrier gas molecules compared to the analyte molecules, almost all ions are initially so-called reactant ions that transfer their charge to the analyte molecules through chemical ionisation due to collisions after the ionisation [138]. When the carrier gas is air, the positive reactant ions are primarily protonated water clusters ($H^+(H_2O)_n$) and the negative reactant ions are deprotonated water clusters ($O_2^-(H_2O)_n$). The role of water as a reactant ion makes humidity an important factor in the measurements, since the amount of reactant ions is directly linked to it. In a complex sample, the analyte molecules compete for the charge transfer, and, in addition to the abundance of a specific analyte, the proton (or electron) affinity of the substance is an integral factor. Proton affinity describes the heat energy associated with the removal of a proton from a molecule, expressed as kJ/mol [143]. The higher the affinity, the more prone the molecule is to ionise through chemical interactions, i.e. collisions with ions with lower affinity, in the ionisation process. If air is used as the carrier gas, molecules with a lower proton affinity than water will not be detected, since the charge transfer in the collisions favours the substance with the highest affinity. In a situation where the abundance of different analyte molecules (with higher affinity than water) is similar, each analyte molecule group can collide with reactant ions (or rarely with each other) and form distinct ion peaks in the output [144]. However, the differentiation of substances diminishes if the sample has a very large concentration of some molecules with high proton affinity. As an example, the proton affinity of water has been shown to increase in relation to humidity as the size of the water molecule clusters increase [145]. In extreme moisture, the proton affinity of water can be so high that the molecules "steal" all the charge and cause a situation in which other analyte molecules stay neutral and do not reach the detector, resulting in only a single signal peak in the spectrum [144,146]. In a simplified setting, in normal humidity (and in the positive polarity) with one measured substance in a carrier gas, a neutral analyte molecule (M) first forms an intermediate adduct ion as a result of a collision with a reactant ion ($H^+(H_2O)_n$). This intermediate product is stabilized by dissociation back to the reactants, or by the displacement of water, which produces a protonated monomer and neutral water (Equation 1). If the sample concentration is high enough, the protonated monomer can be accompanied by another sample neutral molecule to form a proton-bound dimer (Equation 2) [146].

In theory, trimers and tetramers can also be formed, but their lifetime after the ionisation region is too short for them to reach the detector, as they more easily dissociate back to their reactants [138]. An example of how the reactant ion peak (RIP), protonated monomer and proton-bound dimer are seen in the DMS output in a simple measurement of 1-butanol is demonstrated in Figure 7.

$$M + H^+(H_2O)_n \leftrightarrow MH^+(H_2O)_n \leftrightarrow MH^+(H_2O)_{n-x} + xH_2O \quad (1)$$

$$MH^+(H_2O)_n + M \leftrightarrow M_2H^+(H_2O)_{n-x} + xH_2O \quad (2)$$

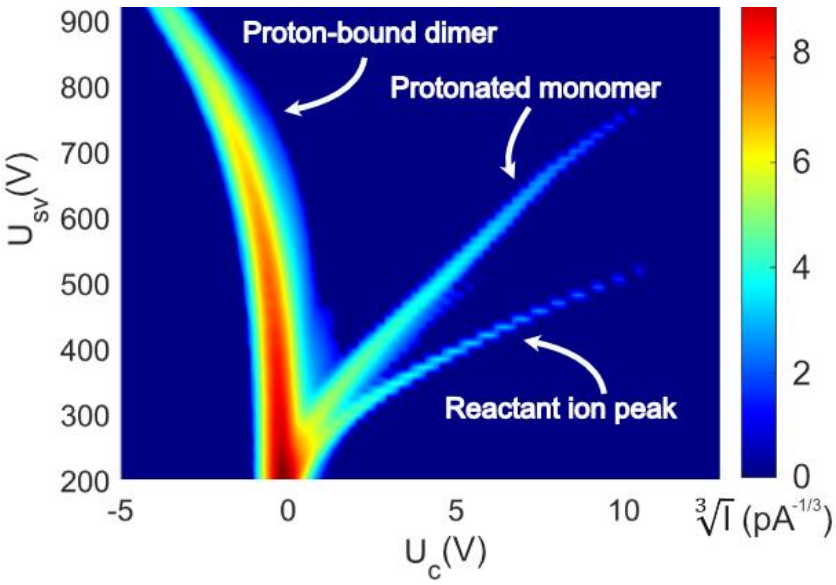


Figure 7. *DMS measurement of 1-butanol at a concentration of 40 parts-per-billion (volume) showing the formation of characteristic ion peaks. The ion peaks are formed with different separation voltage (U_{SV}) and compensation voltage (U_C) pairs and produce a pA-range current that is represented as colours in the plot. In this image, the current is presented as cube root values to enhance peak visibility*

After the analyte ions are formed, they are subjected to an asymmetric radiofrequency electric field in the DMS filter region, where the strength of the field alternates between high (> 20000 V/cm) and low (~ -2000 V/cm) with a duty cycle of 5%–30%, so that the integral is zero [147,148]. Each ion swarm that enters the DMS filter region can be thought to have a certain mobility coefficient (*K*) that depicts their drift velocity (*v*) in an electric field (*E*) [139]:

$$K = \frac{v}{E} \tag{3}$$

The drift velocity of a measured ion swarm depends on the pressure (*P*) and temperature (*T*) inside the DMS filter region, which is why a normalised form of the mobility coefficient is often used, when describing and comparing ion mobilities. This so-called reduced mobility coefficient (K_0) is normalised to a standard pressure of 760 torr and temperature of 273 K as follows:

$$K_0 = K\frac{P}{760}\frac{273}{T} \tag{4}$$

The mobility depends on the collisional cross section and charge of the ion species, and, as such, similar mobility values between different ion species are possible. However, the mobility coefficient does not remain constant, and it changes based on a nonlinear ion-specific characteristic function (*α(E/N)*). This phenomenon is utilised in the rapidly changing electric field of the DMS filter region. The field-dependent mobility (*K(E/N)*) can be equated as:

$$K\left(\frac{E}{N}\right) = K_0\left(1 + \alpha\left(\frac{E}{N}\right)\right) \tag{5}$$

The field dependence is notated as (*E/N*), since the motion of ions is affected by the combination of the strength of the field (*E*) and the neutral molecule density (*N*). The movement of ions is increased along with increasing *E*, but, conversely, the losses in kinetic energy increase with a higher *N* due to increased collisions [149]. The *α(E/N)* function of an ion species is a complex phenomenon, but it can be explained by the clustering of an analyte with carrier gas molecules (or added dopant) [150]. As an example, species with a positive change in the *α(E/N)* function in increasing field strength can form a cluster (e.g., $MH^+(H_2O)_n$) with a certain collisional cross section in a low field. Then, in a high field, as a result of increased kinetic energy and the resulting energetic collisions with neutrals, the vibrational energy of the ion cluster increases, and this disperses the cluster. As a result, the collisional cross section of the analyte ion is reduced, which, in turn, increases its mobility. When the strength of the field is lowered again, the cluster re-forms [151]. A depiction of this phenomenon is shown in Figure 8. As this alternation of a low and high field is repeated, the net movement of the analyte is positive compared to its movement in a standard low field. It is worth noting that the field dependence can also be negative, which holds true especially for proton-bound dimers and large ion species. This is likely because the decrease in the collisional cross section due to declustering is not enough to compensate for the increased number of collisions in a

high field that reduce the kinetic energy of the analyte. Thus, these types of ions actually move faster in low fields [151]. Either way, ion species that could have similar trajectories in a constant field can be separated in DMS due to this phenomenon.

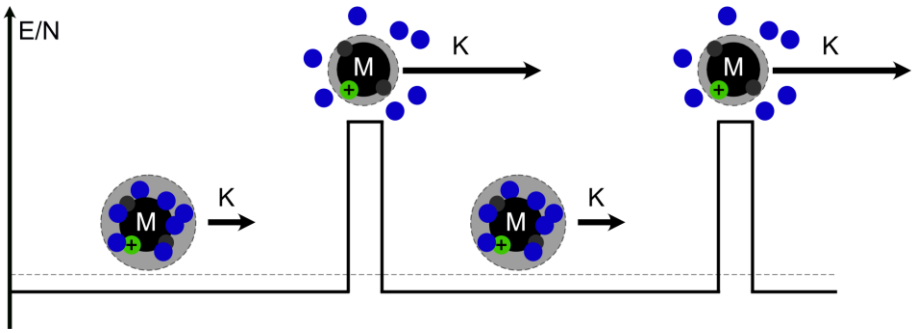


Figure 8. *Depiction of an ion species (M) with positive field-dependent mobility. In a high electric field (E/N), the collisional cross section of the analyte cluster is reduced, and its mobility (K) increases.*

If the DMS filter region would be governed by the alternating electric field alone, only ions that have no field dependence would be able to pass, as their net movement does not change in relation to the field alternation. To selectively measure field-dependent ions as well, a direct current electric field is imposed to the filter region to change the trajectory of those ion species that would otherwise neutralise as they collide with the filter electrodes [152]. A compensation field of only a small fraction of the strength of the alternating field is enough to correct the trajectories of the ion species affected by the field dependence. The strength of the compensation field is dependent on the compensation voltage, U_C (or V_C), and is usually in the range of approximately -15 V to 15 V, compared to the separation field voltage, U_{SV} (or V_{RF}), values of 200 V to 1000 V. By altering the U_C value, ion species that have a positive or negative α(E/N) will reach the detection area, where they will produce a current signal in the range of pA upon impact with a detector plate. In the detection area, an electric field created by positive and negative bias voltages (V_{B+} and V_{B-}, typical values ± 3 V) drive the ions to the detector plates. A schematic representation of the structure of the main entities and operation principle of a DMS instrument is depicted in Figure 9.

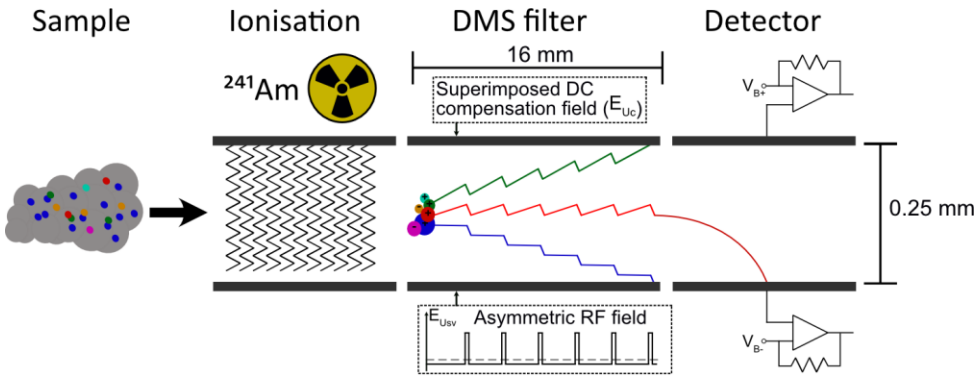


Figure 9. *The operation principle of DMS. The filter dimensions and ionisation source (Americium-241) are based on the DMS used in the thesis research. Only ions with a net mobility of zero will reach the detector area where an electric field created by bias voltages (V_{B+} and V_{B-}) force the ions to a detector plate.*

By systematically measuring different pairs of U_C voltages and the asymmetric electric field maximum voltages (U_{SV}), all ion species of the sample can ideally be measured in one measurement sweep to produce a comprehensive view of the molecular mixture of the sample gas. Typically, this is the way to produce the output of a DMS measurement in the form of a *dispersion plot.* However, the measurement of all U_C-U_{SV} voltage pairs is not ideal, especially if the allotted measurement time is limited. Each measured voltage pair adds tens of milliseconds to the total measurement duration, but some parts of the dispersion plot do not contain information on the measured sample [148]. At low U_{SV} field voltages, the change between a high and low field is not enough to disperse the ion clusters, and, as a result, the only visible signal peak is seen at compensation voltages near zero. On the other hand, at high U_{SV} voltages, the dispersion of the species increases, but, at the same time, the increased movement towards the filter walls and the resulting neutralisation diminish the total yield of ions that manage to pass through the filter intact. This means that the signal-to-noise ratio is much lower with high U_{SV} values. Additionally, most substances do not have a higher positive $\alpha(E/N)$ than water, meaning that the area "beneath" the reactant ion peak is almost always devoid of information (if the carrier gas is air without additional dopants). The same applies to the area located to the left from the characteristic peak for proton-bound dimers and other heavy ion clusters, since substances with a lower negative $\alpha(E/N)$ are non-existent. Figure 10 highlights the information areas of the dispersion plot in an

example measurement of α-pinene, which is a natural hydrocarbon with various pharmacological effects and potential clinical applications [153].

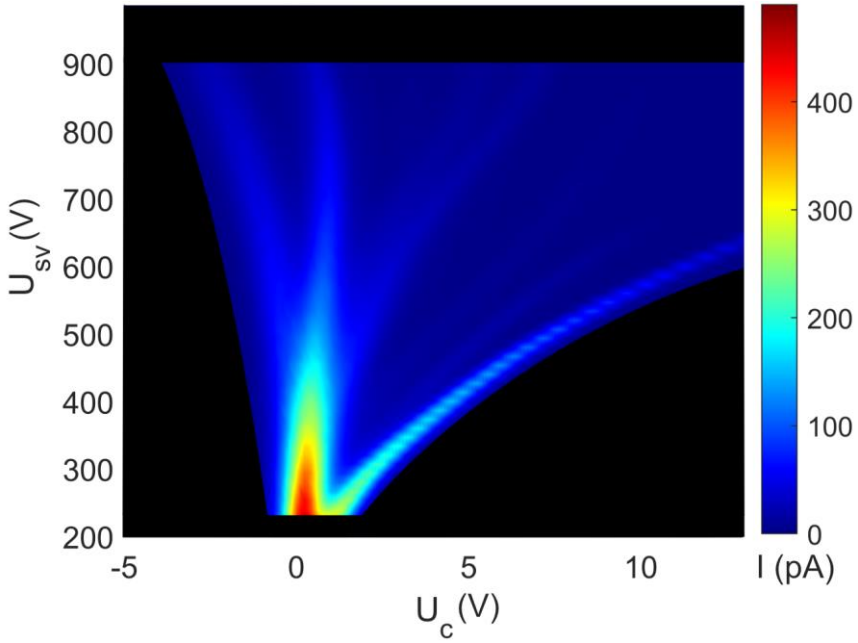


Figure 10. *A dispersion plot of α-pinene with areas of no information darkened. The colours in the plot represent the number of ions that produce the measured pA-range current, with certain compensation voltage (U_C) and separation voltage (U_{SV}) values.*

DMS as a technology is more robust than MS. Individual instruments can retain their functionality without major maintenance or calibration for several years, which is why the technology has been used for air quality monitoring in space stations and submarines [154]. In biomedicine, DMS has been utilised in, for example, the identification of bacteria and in breath analysis [14,155,156]. DMS has also been shown to be applicable to quantifying a mixture of phospholipids, lecithin, in controlled measurement conditions [157]. Still, the translation of a DMS spectrum into definitive molecules or compounds gets increasingly complicated in heteromolecular media due to the increased degrees of freedom through molecular interactions that affect the field dependence of the ion swarm and that are hard to model or simulate [158,159]. Thus, when considering the biomolecules that are at the forefront in tissue identification, there has not been extensive research in the use of DMS alone in proteomics or metabolomics. However, DMS has been used together with MS in these applications [160,161]. Because DMS can be thought of as a filter

that only allows ion swarms with a certain electric field dependence ($\alpha(E/N)$) to pass, the separation capability of MS systems can be greatly increased if the technologies are used in tandem. In addition, the identification of isomers and other molecules with a similar mass but a different structure is possible, if DMS is used prior to MS. The added detection capability functions well in omics applications, but clinical applications would benefit greatly from the exemption of a mass analyser due to the reduced cost, size and complexity. However, since DMS is not equipped for the analysis of specific molecules similarly to MS, if the application is solely dependent on the output of a DMS device, the data analysis methods to be applied require careful consideration.

2.5 Data analysis

The DMS output spectrum is a representation of the measured gas sample, but especially in a heterogeneous sample matrix, the subtle differences in the spectra are not always clear upon visual inspection. Therefore, identification through machine learning methods from the values of the DMS output data can be a more viable option. Machine learning is used to find generalizable patterns from (especially multidimensional) data and utilising these patterns to predict the type or class of new data [162]. In the studies presented in this thesis, the analysed data entailed the DMS dispersion plots that were produced for each sample. This means that the measured current signal with each U_C-U_{SV} pair was treated as a separate input variable in the data analysis and classification.

Data can be separated into different groups using labelled or unlabelled data. If unlabelled data are used, the aim is to find features that create separability within the dataset without specific training, reinforcement or information regarding the true class or type of the samples [163]. Clustering methods in which samples are grouped together based on similarity of the data features are an example of this type of *unsupervised* learning. In the studies presented in this thesis, unsupervised learning was not extensively utilised. Still, it could be useful in DMS-based tissue imaging in pathological applications to highlight grossly distinct areas based on their smoke profile, without training a model or labelling, as demonstrated in Publication III. However, to assign a class and explicitly predict the type of operated tissue, *supervised* learning is needed. In supervised learning, a classification (or a regression) model is trained with example data to learn a function that transforms numerical input data to a categorical output [164].

An example of a supervised classification method is linear discriminant analysis (LDA), which was utilised in regularised form in all publications presented in this thesis. In LDA, two or more known groups are separated by a linear decision boundary (or boundaries) in such a way that the variation within the group is minimised and the between-group variance is maximised [165]. Mathematically expressed, the principle of LDA for a two-class problem is to find a projection vector, $\vec{w}$, to maximise the value of a criterion function, $J(\vec{w})$:

$$J(\vec{w}) = \frac{\vec{w}^t S_B \vec{w}}{\vec{w}^t S_W \vec{w}}, \tag{6}$$

where S_B represents the between-class scatter matrix and S_W represents the within-class scatter matrix of the data. The solution vector that optimizes the function is called the *Fisher's linear discriminant* and can be expressed as:

$$\vec{w} = S_W^{-1}(m_1 - m_2), \tag{7}$$

where m_1 and m_2 are the sample means for the separable classes. The Fisher's linear discriminant thus presents a linear function that maximises the distance between class means and minimises the variance within the classes. In LDA, the additional assumption is made that the data are normally distributed and that the classes share a common covariance matrix. For a simple setting of two-dimensional data and two separable classes, the decision boundary for class separation is a normal of the Fisher's linear discriminant to which the projection of the data points produces maximum separability. A visual depiction of this situation is presented in Figure 11. If the data has three dimensions or has been reduced to such, the decision boundary is a two-dimensional plane [166]. For higher-dimensional data, such as the DMS data in this thesis, the decision boundaries are also higher-dimensional planes, and, as such, cannot be easily represented visually.

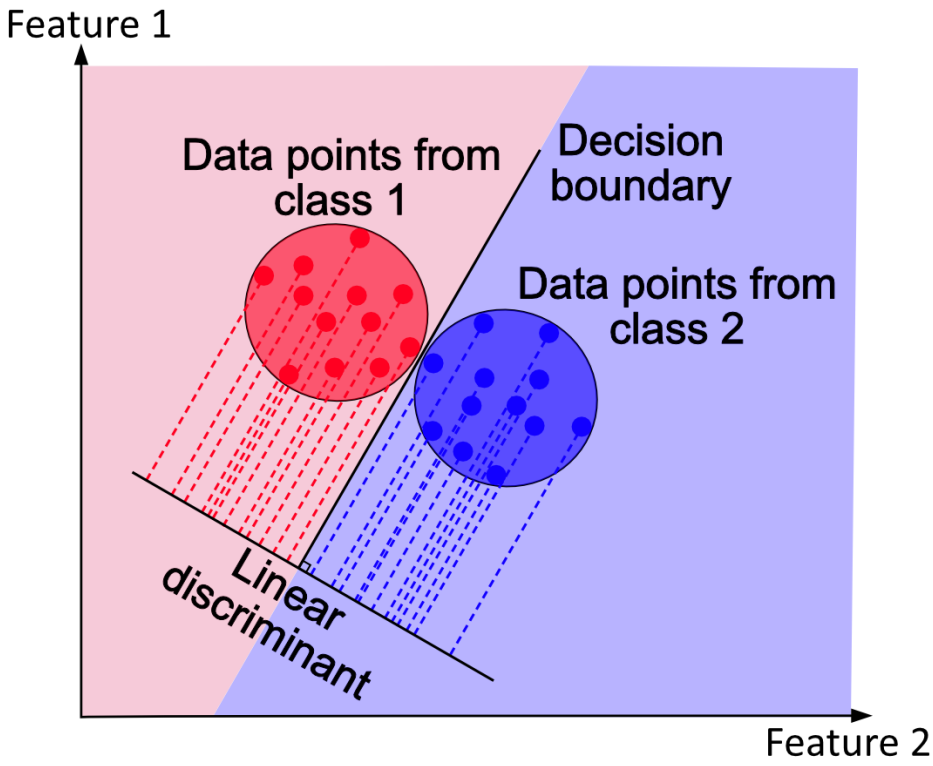


Figure 11. *Simplified illustration of linear discriminant analysis in a two-dimensional feature space.*

It can be beneficial to reduce data dimensionality to highlight informative areas in the feature space and to improve the classification performance by removing redundant or noisy features that contain no information and add no value to class separation [167]. To reduce the data dimensions, a process called forward sequential feature selection (FSFS) was utilised in Publications II, III and IV. In FSFS, the features of the data are selected one by one in the order of increasing classification accuracy [168]. In practice, this means that the feature selection algorithm explores the emptied feature space and first selects the feature that produces the highest classification accuracy for a given model. After the first selection, another feature is selected with the same principle. This process is continued until a stopping criterion is met, which can be defined as a specific number of features selected or a point where significant improvement in classification is not achieved with further feature additions [169].

In addition to LDA and FSFS, in Publication IV, convolutional neural network (CNN) analysis was experimented. Since then, CNN has also been used in the prediction of lecithin concentration from DMS data [157]. CNN is a more complex supervised learning method that is especially utilised in image classification [170]. Since the output of DMS measurements can be presented as a two-dimensional dispersion matrix that can be interpreted as an image with distinct peak shapes and intensities, CNN provides a feasible alternative to the analysis of DMS data. In principle, CNN is a specialised case of an artificial neural network that loosely mimics the function of the human visual cortex, with large numbers of interconnected neuron layers that transform the input data to a categorical output by assigning and adjusting weighing coefficients for the features of data based on a training set. Before the neuron layers, however, the data in CNN is first passed through convolutional layers that basically function as a filter for the raw (visual) data to extract high-level two-dimensional features (e.g., peak edges) [171]. Due to its complexity and extensiveness of parametrisation possibilities, the training of CNN models can be computationally demanding and slow, and it often requires a large amount of training data. However, rather than treating each feature as independent, CNN considers the continuous structure of the data and can be used to model nonlinear phenomena. Thus, with a sufficient amount of training data, CNN is likely to outperform the simpler methods in terms of classification performance, especially in complex classification problems [157,172].

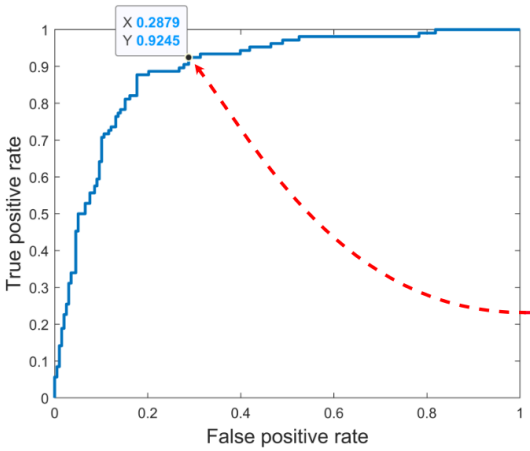
The classification performance of a machine learning model is traditionally measured with classification accuracy, which expresses the percentage of correctly classified samples of a test set or sets. In a diagnostic setting with only two separable classes, sensitivity and specificity are also often used as performance metrics. Sensitivity describes the fraction of positive samples that have been correctly identified, whilst specificity does the same for the negative samples [173]. In addition, positive predictive value (PPV) and negative predictive value (NPV) are used as metrics to assess the value of the corresponding test result to a clinician, since they account for the prevalence of the tested condition [174]. A more detailed representation of the classification performance can be given in the form of a so-called *confusion matrix*. A confusion matrix presents the classification results in two dimensions, so that the true and predicted classes are presented by table rows and columns [175]. Table 3 shows an example confusion matrix for a diagnostic setting, with the formulas for the diagnostic metrics also provided.

Table 3. *An example confusion matrix for a diagnostic setting for the identification of malignant samples with formulas for sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV).*

		Predicted class		
		Malignant	**Benign**	
Actual class	**Malignant**	**98** **True positives (TP)**	**2** **False negatives (FN)**	**Sensitivity = TP/(TP+FN) = 98%**
	Benign	**6** **False positives (FP)**	**94** **True negatives (TN)**	**Specificity = TN/(TN+FP) = 94%**
		PPV = TP/(TP+FP) ≈ 94%	**NPV = TN/(TN+FN) ≈ 98%**	

The performance of a classification model in a binary setting can also be examined with a so-called receiver operating characteristic (ROC) analysis, as in Publication II. ROC depicts the relationship between the specificity and sensitivity of a model by plotting the *true positive rate*, which is an another term for sensitivity, as a function of the *false positive rate* [176]. The false positive rate portrays the probability of a false positive and its value can be written as 1-true negative rate (i.e. 1-specificity). Classification models assign a class to a new sample based on the features learned from the training data. Based on the similarity of the sample compared to the training data, classification models also produce a posterior probability score for the prediction, which basically shows the perceived certainty (probability) of the sample belonging to a certain class. The ROC curve can be produced based on these scores. In a standard situation where a sample is equally likely to belong to either the positive or the negative class, the threshold level for the posterior probability score for assigning a class is 0.5. As an example, if the classification model produces a score of 0.57 for the prediction of malignant tissue, the displayed classification result is malignant. However, the threshold level can be adjusted if, for example, the sensitivity of the diagnosis is deemed more important than the specificity. This can be of value, when the class sizes are significantly unequal. In ROC analysis, the resulting true positive and false positive rates with different threshold levels are plotted as a curve and often accompanied by the value for the area under the curve (AUC) to give a numerical value for the performance [177]. In the ideal case of a

perfect classification model that detects every positive (and negative) sample without fail, the ROC curve will form a rectangle with an area of 1, since all predictions will be correct regardless of the threshold selection. If each negative sample receives a higher posterior probability score than the positive samples, then every sample will be misclassified and the AUC will be zero. An example ROC curve drawn from the breast cancer data used in Publication II is presented in Figure 12, alongside the values for the posterior scores and their corresponding metrics.



True class	Predicted class	Posterior probability score	True positive rate	False positive rate
⋮	⋮	⋮	⋮	⋮
Benign	Benign	0.2981	0.8868	0.2525
Benign	Benign	0.2883	0.8868	0.2576
Benign	Benign	0.2836	0.8868	0.2626
Benign	Benign	0.2831	0.8868	0.2677
Malignant	Benign	0.2827	0.8962	0.2677
Benign	Benign	0.2801	0.8962	0.2727
Benign	Benign	0.2722	0.8962	0.2778
Malignant	Benign	0.2618	0.9057	0.2778
Benign	Benign	0.2599	0.9057	0.2828
Benign	Benign	0.2531	0.9057	0.2879
Malignant	Benign	0.2459	0.9151	0.2879
Malignant	Benign	0.2437	0.9245	0.2879
Benign	Benign	0.2428	0.9245	0.2929
Benign	Benign	0.2414	0.9245	0.2980
Benign	Benign	0.2408	0.9245	0.3030
Benign	Benign	0.2314	0.9245	0.3081
Benign	Benign	0.2305	0.9245	0.3131
⋮	⋮	⋮	⋮	⋮

Figure 12. *An example receiver operating characteristic (ROC) curve for DMS-based identification of breast cancer. If the threshold for assigning the positive class to a sample based on the posterior probability score would be below 0.2437 and above 0.2428, as indicated with the red line, all predictions above the line would be malignant. This would increase the sensitivity to 92.45% but decrease the specificity to 71.21%. The ROC curve shows all possible discretised results of the threshold variation.*

When presenting any of the metrics for classification performance in a supervised model, it is crucial to disclose the details of the datasets for which the results hold true. The importance of transparency is due to a phenomenon called *overfitting*. Since a supervised model is trained with a certain dataset, the predictions it performs are based on the exact properties of that particular dataset, even if some of those properties are not descriptive of the true class differences but only exhibit noise or otherwise random variation [178]. Therefore, if the same dataset would be used as an input for the classification, the result could be 100% correct. However, this says nothing of the model's capability to predict new data, which is why a process known as *cross-validation* is often used.

In cross-validation, part of the same dataset is excluded from the training process and used as a test dataset containing data points that are essentially new to the model. Commonly, cross-validation is performed by splitting the dataset multiple times with different (yet equal-sized) sample partitions for training and testing, which are done separately for each split. The classification performance is then reported as the average of the classification with each partition to give an estimate of the performance for the classification problem [179]. Often, the dataset is randomly partitioned into k folds, where k is a natural number – for example, 5 or 10. A special case of k-fold cross-validation is leave-one-out cross-validation (LOOCV), where each sample is used as a test set in turn. LOOCV is computationally more demanding, since the training phase must be carried out as many times as there are samples in the dataset, but the method can be useful if the dataset is limited in size and it is not possible to gather more data to use as a separate test set. However, the problem with reporting cross-validated results lies in the fact that the data within a certain set is rarely independent. For example, a certain set of data could be measured on the same day in similar environmental conditions, samples can come from the same source (e.g., patient), and/or the measurement instrument could have some contamination or baseline signal during the measurements.

All of these factors could affect the measurement result. Therefore, a model can be overfitted to features of the training data that do not, in reality, describe the differences between the classes and still yield good results with random split cross-validation. With fully independent testing datasets, overfitting can be more effectively distinguished, and the reported result will reflect the generalisability of the performance better. As such, completely separate independent test data are always preferred and should be standard practice, when presenting classification results [180,181]. However, especially with small sample sizes, LOOCV or k-fold cross-validation are also sometimes used, even though they are susceptible to producing over-optimistic results. Examples of this can be found in some of the studies that report the tissue identification performance of current margin assessment technologies [106,109,125,182]. The same applies for the studies presented in this thesis, which is why overfitting and its proper mitigation are aspects that could have been improved upon in the study design of the research in this thesis.

3 MATERIALS AND METHODS

The focus of this thesis is the utilisation and development of the DMS-based tissue analysis system from preliminary proof-of-concept animal studies to the first pilot study in an operating room. During each phase, the system was gradually improving to facilitate the real-time analysis of tissues *in vivo*. This chapter explains the technical evolution of the system and the research setup for each publication. The main aspects of the research setup in each phase are listed in Table 4 at the end of this chapter.

3.1 Porcine tissue identification

After conceptualising the idea of a DMS-based tissue analysis system, the aim of the first research setup was to provide proof that the concept works. For this purpose, ten landrace pig tissue types were obtained from a slaughterhouse (Paijan tilateurastamo, Urjala, Finland). All tissues, except muscle and fat from flank pieces, were offal and would have been disposed of as waste. The tissues were kept in a freezer at -18 °C and thawed before the measurements. The choice to use porcine tissue in the feasibility testing of the system was mainly based on the availability of the material and ethical considerations. However, porcine tissues are also a relatively good substitute for human samples, due to their similar anatomy, physiology and metabolism [183,184]. Presumably, these similarities also extend to their properties regarding tissue evaporation and coagulation.

The tissues were measured in three phases. In the first phase, ten samples from five tissue types (lung, liver, fat, skeletal muscle and renal cortex) were measured during one measurement session. In the second phase, five more tissues (cardiac muscle, tongue, skin as well as grey and white matter of the brain) were added, and each tissue was sampled 60 times, resulting in 600 measurements. In the third phase, the remaining tissue pieces from the renal cortex, skeletal muscle and fat were measured during one session to form a small testing set with 10 samples from each tissue. The tissues in the first study were measured with the first iteration of the system. In its infancy, the system was only the sum of its parts without much

consideration for the appearance or optimal function. Nonetheless, all the parts of the first system functioned as the core in future system iterations as well. The measurement system can be divided into three parts: tissue sampling, surgical smoke sample pre-processing and DMS measurement.

The tissue was sampled by an electrosurgical instrument (blade: HF 9805-24, HEBUmedical GmbH, Germany; power source: Itkacut 350 MB, Innokas Medical Oy, Finland) that was attached to a modified 3D printer platform (REPRAP, Mendel Prusa i3 kit, KitPrinter3D, Spain), where the filament extruder was replaced with the electrosurgical blade. This allowed XYZ-directional control of the surgical instrument. The sampling was controlled by a graphical user interface (GUI) coded with MATLAB (The MathWorks Inc., USA) that allowed for the creation of automated incision sequences with fixed cut durations and depths. The sampling stage created the tissue smoke that was then filtered and diluted in the sample pre-processing stage. The filter removed large contaminating particulate matter from the smoke sample and the dilution decreased the sample concentration in order to prevent measurement saturation. After the sample pre-processing, the smoke was measured with the DMS (ENVI-AMC®, Environics Oy, Finland). Together, the sampling stage, pre-processing stage and the DMS formed the automatic tissue analysis system (ATAS). The first iteration of the ATAS is illustrated in Figure 13.

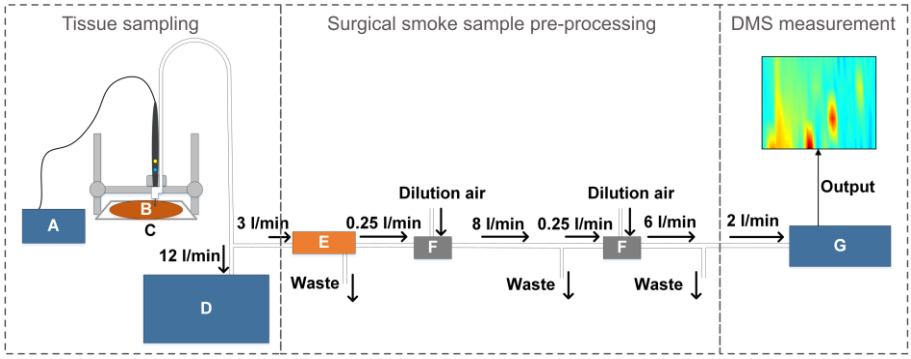


Figure 13. *The first iteration of the automatic tissue analysis system. A) Electrosurgical unit. B) Tissue sample. C) Sampling platform. D) Surgical evacuator. E) Corona discharge filter. F) Dilution system using purified and dried dilution air. G) ENVI-AMC® differential mobility spectrometer. Modified and reprinted with permission from Springer Nature: Springer Nature, Annals of Biomedical Engineering, Tissue Identification in a Porcine Model by Differential Ion Mobility Spectrometry Analysis of Surgical Smoke, Anton Kontunen et al., Copyright © 2018, Biomedical Engineering Society.*

Most parts of the first ATAS system were commercial devices chained together, but the filter was custom-built with the requirements of a real-time tissue analysis system already in mind. The simplest and most inexpensive solution for removing the large particulate matter could have been to use glass fibre filters, but the resulting particulate residue in the filters and pressure differences would likely have caused issues with carry-over residue and sample flow dynamics [185]. Thus, an electric filter based on a corona discharge was incorporated in the system. The electric filter consisted of a needle electrode in high potential (~5 kV) and a grounded copper casing. The principle of operation relied on the ionisation of the initial sample stream, after which the ionised particles would migrate towards the grounded casing. After neutralisation, large particulate matter would travel to a waste outlet while smaller molecules would enter a sample outlet located at the root of the needle electrode. The filter was conceptualised, simulated and tested before the research presented in this doctoral thesis, and measurements conducted for a master's thesis work revealed that the filtration efficiency was at least 99.99% for particulate matter with a diameter of 7 nm to 10 µm [186]. This infers that the harmful airborne substances, such as potential soot particles, bacteria and viruses, are extremely effectively removed from the measured sample [187]. Therefore, the electric filter was considered an important aspect of the measurement system to prevent contamination, and the core design was maintained in all iterations of the tissue analysis system.

Another integral part of the system that would remain unchanged throughout the thesis research is the ENVI-AMC DMS device. Its DMS filter channel is 0.25 mm high, and the device can create a 250 kHz separation field with strengths of 10–80 Townsend (Td). Td is commonly used when discussing DMS, since the unit also takes into account the number of neutral molecules in the electric field, i.e. E/N [188]. One Townsend equals 10^{-17} V/cm^2, so the field strengths of 10–80 Td in standard conditions (20 °C, atmospheric pressure) and with, for example, 2.5×10^{19} neutral molecules in a cubic centimetre can thus be written as 2500–20,000 V/cm. The compensation field produced by the ENVI-AMC device can be between -1.5 Td to 1.5 Td. In the device user interface, the field strengths are controlled by setting a starting and stopping value for the compensation and separation field voltages. In addition, the number of increments for both fields is selected. This results in one full measurement containing a selected number of successively measured U_C-U_{SV} voltage pairs that together form the dispersion matrix.

In the first study, the compensation voltage was incrementally increased from -0.8 V to 9 V in 40 steps and the separation voltage from 440 V to 770 V in 4 steps. This resulted in a relatively small dispersion matrix of 160 values (per ion polarity). The voltage limits were chosen based on the assumption that those areas in the dispersion plot were most likely to contain the relevant information as described in Section 2.3. The low resolution was due to the required, relatively short sampling time. The electrosurgical smoke was produced for approximately five seconds per cut, so the measurement duration had to be limited. Longer cuts were not considered, since the amount of surgical smoke with the five-second automated cut was already substantial enough to cause a significant carry-over signal that required a three-minute waiting period between the measurements. In the first study without prior information on the classification performance, the decision was made to include the measurement of both polarities, even though this increased the measurement time to a total of 12 seconds, since both polarities could not be measured simultaneously with the ENVI-AMC.

The 160 values from both polarities were used as the data features to classify the tissue types. The classification models were created with MATLAB using the Statistics and Machine Learning Toolbox function for LDA (*fitcdiscr*) with regularisation. In LDA analysis, the assumption is made that the discriminated populations share a common covariance matrix. The covariance matrix can be modified with a regularisation term, so that the sample-based estimate of the covariance matrix is replaced with an estimate that lessens the importance of individual features [189]. Regularisation is often mandatory for DMS data because if each value in the spectra is used as a predictor in a model, the number of dimensions can, in some cases, exceed the number of observations, and, in addition, the features can be correlated or completely non-informative. This means that the prerequisite of a common covariance matrix is not achieved and non-regularised LDA cannot be used [190]. The decision to use (regularised) LDA was based on exploratory analysis of preliminary DMS data where it produced the highest classification accuracies in tissue identification, when compared to other supervised classification methods, such as k-nearest neighbours and support vector machines. LDA has also been successfully utilised with IMS and DMS data for the detection of prostate cancer and gastrointestinal diseases from urine samples [191,192]. In addition, LDA has been extensively utilised in REIMS studies and shown to outperform several other classification methods in discriminating cancer from benign samples based on MS data from blood serum [193–195]. The performance of the LDA classification models used in the different phases of the measurements in Publication I were evaluated with

LOOCV or 10-fold cross-validation, or by holding out 30% of the data as test samples.

3.2 Breast cancer identification

The next step in the research was to test the system with clinically relevant tissues. Since breast cancer surgery is one of the main application areas for the proposed margin assessment system, human breast samples were chosen as the material for the proof-of-concept study. With the approval of the Ethical Committee of Tampere University Hospital (code R17007), refrigerator-stored 4 mm punch biopsy samples of breast cancer and benign tissue collected from a total of 21 patients were used as sample material. The total number of analysed measurements was 304 (106 malignant, 198 benign). Since the biopsy samples were much smaller in size than the tissue pieces used in the first study, the sampling process needed to be refined. Instead of placing the samples directly on the sampling platform and producing line incisions to create the tissue smoke, an aluminium well-plate that would help to position the samples for more controlled and sparing point-like sampling was used. The well plate kept the small samples in place and enabled the creation of an automatic measurement sequence with pre-determined distances between the samples.

Other parts of the system used in the breast cancer study had some alterations compared to the first study as well. More specifically, the pre-processing stage of ATAS was modified by adding a stainless-steel sample chamber for the filtered and diluted smoke. The chamber allowed for a longer measurement of the sample through a controlled release of smoke with proportional valves. This enabled the acquisition of higher-resolution DMS spectra from the tissue samples. The aim was to highlight the areas of interest in the spectra between malignant and benign samples more accurately. Furthermore, after the initial study with porcine tissues, the negative dispersion plots were not utilised, since the acquisition of the negative spectra with the DMS instrument would have doubled the measurement time. Even though this meant that potential identifiable molecular features could be lost, the classification performance with only the positive spectra was shown to be similar to the performance with both polarities in previous measurements with porcine tissues. This suggested that, at least in the porcine tissue identification, the differentiable features that manifest in the positive spectra are sufficient for achieving good classification performance. In addition to the inclusion of the sample chamber and exclusion of the negative spectra, the pneumatic lines of the system were shortened to reduce the

carry-over effect by limiting the available adhesion area for the gas molecules. The sample pre-processing unit was also packaged into a box, rather than being a scattered set of pressure ejectors, valves and tubing as before. An illustration of the ATAS used in Publication II is presented in Figure 14. The system used in the breast cancer identification was also utilised to identify different types of brain cancers successfully, as reported in a separate publication [196].

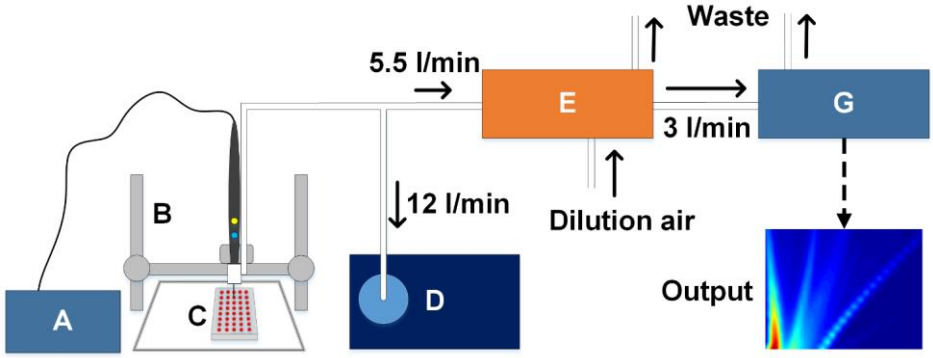


Figure 14. *The second iteration of the automatic tissue analysis system. A) Electrosurgical unit. B) Sampling platform. C) Sample well plate. D) Surgical evacuator. E) Sample pre-processing unit. G) ENVI-AMC® differential mobility spectrometer. Reprinted with permission from Elsevier: Elsevier, European Journal of Surgical Oncology, Identification of breast tumors from diathermy smoke by differential ion mobility spectrometry, Maiju Sutinen et al., Copyright © 2018 Elsevier Ltd, BASO ~ The Association for Cancer Surgery, and the European Society of Surgical Oncology. All rights reserved.*

The modified ATAS enabled the acquisition of a dispersion plot with 1620 U_C-U_{SV} pairs for each measured sample. The limits of the compensation and separation voltages were kept the same as in Publication I (-0.8–9 V and 440–770 V, respectively), but the number of steps was increased to 90 for compensation voltage and 18 for separation voltage. The DMS measurement duration was approximately one minute, after which the sample chamber was cleansed with purified air for an additional minute. The data produced in the measurements were again classified by first using LDA with 10-fold cross-validation and then separately by training a model with 70% of the data and holding out 30% as test samples. However, as an addition to the data analysis, FSFS was also implemented in MATLAB to study how many features of the spectra are needed to identify the tissue types. The feature selection was repeated 1000 times with different cross-validation partitions to ensure that the

features selected in the process were not a result of random noise in the features. A ROC curve was also plotted to further illustrate the diagnostic performance of the LDA classification with the features selected with FSFS.

3.3 Tissue imaging

After the proof-of-concept study for *ex vivo* tissue identification by means of DMS analysis of surgical smoke with breast cancer tissue was completed, the aim was to expand the application field of the tissue analysis system and test its function in tissue imaging. This meant that the focus of the development would not be in improving the system towards surgical use. Instead, the focus was on proving that DMS-based tissue identification could also work in pathological applications. Molecular tissue imaging is a field that is dominated by MS analysis [197–199]. However, as with the surgical application, DMS could provide a more inexpensive, more maintainable and simpler solution than MS. The aim in DMS-based tissue imaging would specifically be to complement traditional histopathological analysis by providing an overview of the tissue and to pinpoint potential areas of malignancy that would then be subjected to further analysis.

Since the acquisition of full pathological slides of clinically relevant tissues was not feasible for a proof-of-concept application of an experimental method, the decision was made to return to slaughterhouse offal of porcine tissue for the measurements. This time, the tissue types were lung, renal cortex, renal pelvis, liver, grey matter, white matter, skeletal muscle and adipose tissue. The tissues were frozen and prepared into 3 mm thick slices prior to the measurements. Compared to the first study, the measurement protocol was more refined. Overall, the number of measurements was higher, and the order of measured tissues was randomised to the extent that was possible. However, since the main aim of the study was to image the tissue slices, consecutive measurements of the same tissue slice was a practical necessity. The imaging was conducted by measuring a sequence of 100–400 point incisions (depth 3 mm) per tissue slice. The total number of analysed measurements after the exclusion of saturated measurements and uncertain annotations was 3418.

The measurement system used in the study was similar in function to the ATAS in the first study, but the tubing lengths were minimized by placing the sample pre-processing unit directly on top of the ENVI-AMC and thus providing a direct sample line from the filter to the DMS. The sample chamber from Publication II was removed. This was due to the time constraints of producing the tissue images with

hundreds of cuts. If each measurement were made every two minutes as in Publication II, the applicability of the system in pathological analysis would be difficult to argue, since a comprehensive tissue image would take hours. Furthermore, even though the high resolution likely works best in the exploratory analysis of tissue-specific features in the DMS spectra, the higher resolution also adds to the number noisy and redundant features, since the ENVI-AMC is not capable of a selective measurement of U_C-U_{SV} pairs. Therefore, the resolution was lowered by reducing the steps for compensation voltage to 25 and for separation voltage to 8, resulting in a dispersion plot of 200 values. The lower limit of the separation voltage was set to 340 V instead of 440 V, since FSFS analysis of previous samples indicated that informative areas regarding tissue classification could also be found with lower separation voltages. In addition, the upper limit of the compensation voltage was lowered from 9 V to 5 V, since analysis of previous measurements indicated that the informative value was scarce in the higher U_C-values. With this parametrisation, one DMS measurement took approximately 6 seconds, after which a waiting time of 10–30 seconds was implemented to reduce the carry-over signal. The waiting times were based on the degree of contamination that was typical of different types of tissue based on previous particulate measurements [132]. One minor change was also made to the electrosurgical blade by sharpening it to a needle-like shape to enable more accurate point sampling. The ATAS used in Publication III is depicted in Figure 15.

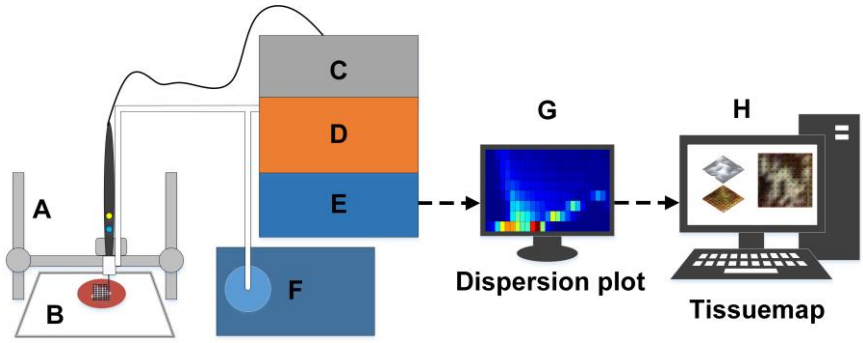


Figure 15. *The third iteration of the automatic tissue analysis system. A) Sampling platform. B) Tissue slice. C) Electrosurgical unit. D) Sample pre-processing unit. E) ENVI-AMC differential mobility spectrometer (DMS). F) Surgical vacuum. G) Monitor showing the dispersion plot of a sample point. H) Computer for data-analysis and visualization of the DMS data. Reprinted with permission from Elsevier: Elsevier, Experimental and Molecular Pathology, Differential mobility spectrometry imaging for pathological applications, Kontunen et al., Copyright © 2020 Elsevier Inc. All rights reserved.*

Even though the system itself was not significantly altered in relation to the previous studies, the data analysis was expanded upon with the introduction of K-means clustering. K-means clustering is an unsupervised method that aims to separate the measurement data into K groups based on how close the features are in relation to each other, without prior knowledge of the classes or their number of measurements [200]. In addition to the unsupervised clustering, another addition to the data analysis was the utilisation of the posterior probability scores of the LDA classification in the visualisation of the results in the form of tissue maps. In practice, this meant that a measured sample slice containing two macroscopically distinct tissue types was photographed and the photograph was overlaid with a heatmap representation of the probability of each sample point belonging to a certain class. The measured samples were classified similarly to the previous studies by using 10-fold cross-validated LDA models. Before the classification, the data was offset-corrected and transformed with the common logarithm to even out the signal intensities in high and low U_{SV} values. The classification was first performed for macroscopically heterogeneous tissues in a binary setting to test the tissue imaging capability, after which all samples were grouped and classified to provide a general overview of the classification performance of multiple tissues. Finally, FSFS was carried out 1000 times for each tissue type in a binary setting (target tissue vs other tissues) to study the tissue-specific features in the DMS spectra.

3.4 Real-time tissue identification

The final intermediate phase before the operating room pilot was aimed to improve the speed of the analysis and to test the tissue discrimination performance of the system without standardised cuts. This entailed that ATAS would be replaced with human-controlled sampling. The sampling was designed to resemble actual surgery with 1–3-second cuts in either skeletal muscle or adipose tissue in random order. The tissues were frozen and thawed prior to the measurements. The waiting time between the cuts was kept at approximately 10 seconds. The goal was to receive representative smoke samples from both tissues rather than to minimise the waiting time. In total, 1059 samples were measured over a period of six measurement days by three different people. An additional validation dataset of 100 samples was also measured 4 months after the initial set. Between the sets, there was a blockage inside the pneumatics of the DMS, which required the disassembly and reassembly of the DMS core.

In general, the measurement system used in the study evolved from the previous versions. The sample pre-processing unit was modified most significantly compared to the previous studies to accommodate the real-time analysis. First, heating of the filter and the sample lines inside the unit was implemented by placing the parts in a wooden box that was heated to 70 °C from the inside by a resistive heating element. Elevated temperature was deemed necessary to minimise the carry-over signal from successive measurements. In a separate study, it was shown that the recovery time of a pneumatic system is proportional to the temperature of the sample lines [201]. The length of the sample lines was minimised even further than previously to reduce the available adsorption area.

Additionally, the electric filter was modified from a passive element to a functional part of the dilution. This was achieved by adding measurement electrodes inside the copper casing to produce a current signal which was proportional to the smoke intake due to the impact of particulate matter ionised by the corona discharge. Based on the current signal, an electric proportional valve that controlled the amount of clean air was adjusted with a proportional-integral-derivative (PID) feedback loop. PID is a common control method that aims to stabilise a system to a target value with minimal overshoot, delay and oscillation, using three types of controllers [202]. In practice, this meant that the current measurement from the corona discharge filter had a baseline value that functioned as the target for the PID control. After sample introduction, the current value changed, which caused the proportional valve to let more dilution air into the filter in order to return the filter to its baseline state. The aim of the dilution control was to keep the smoke sample concentration as constant as possible, despite the variations in tissue material and human-controlled sampling.

Another addition to the system was the measurement of the electrosurgical current with a plastic-encased inductive coil. The induced current created by the electrosurgical cuts was used as a trigger signal for the DMS measurement. The values for the DMS measurement were kept the same as in Publication III, which is to say that the compensation voltage sweep was set to -0.8–5 V in 25 steps and the separation voltage sweep to 340–770 V in 8 steps. In this study, the DMS data were also analysed directly after the measurements. The on-line analysis was performed for the validation set. The measured DMS data were stored in a cloud storage and analysed with a Raspberry Pi computer (Raspberry Pi Foundation, UK) that was connected to an external monitor. In the computer, an LDA classification model trained with the 1059 samples of the first dataset assigned a class to a new measurement from the validation dataset and presented the classification result

alongside the posterior probability for the classification to the user. The whole process from the activation of the electrosurgical instrument to the presentation of the result lasted approximately 9 seconds. The real time tissue analysis system is presented in Figure 16.

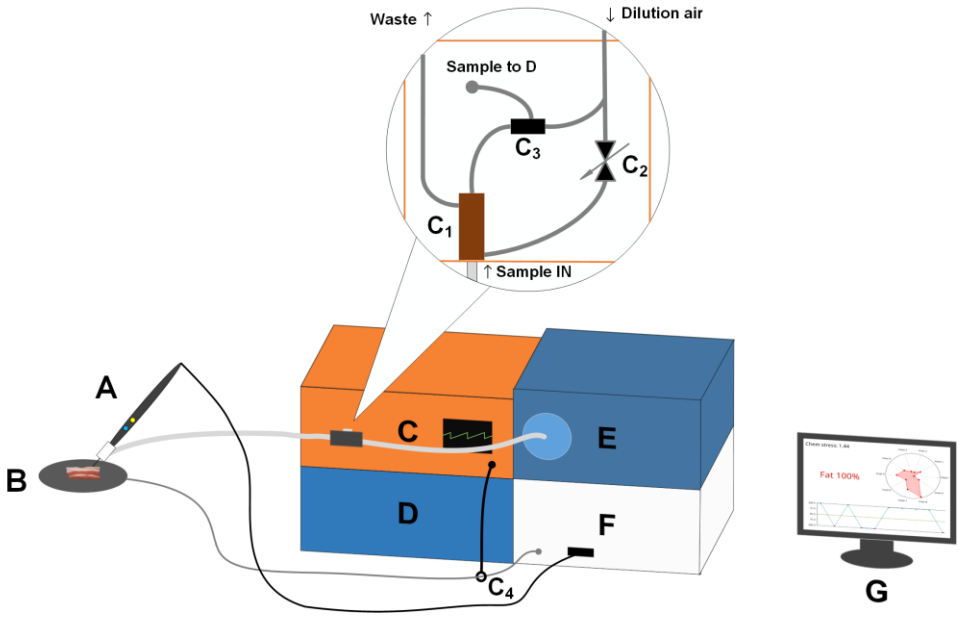


Figure 16. *The fourth iteration of the tissue analysis system capable of real-time analysis. A) Diathermy instrument with the attached 3-metre-long suction tube. B) Tissue specimen on top of the dispersive electrode. C) Sample pre-processing unit with the corona discharge filter (C1), PID-controlled proportional valve (C2), ejector dilution (C3), and an inductive coil to detect the diathermy current (C4). D) ENVI-AMC DMS sensor. E) Smoke evacuator. F) Diathermy unit. G) Surgical user interface. Reprinted with permission from IEEE: IEEE Sensors Journal, Real Time Tissue Identification From Diathermy Smoke by Differential Mobility Spectrometry, Kontunen et al., Copyright © 2021, IEEE.*

In terms of data analysis, this was the first study in which the LDA analysis was complemented with another supervised learning method, CNN. In addition, the data were offset-corrected and normalised. The initial set of 1059 samples were classified with a 10-fold cross-validated LDA and CNN models. The validation set was then classified using the initial set as training data with both methods. However, due to the poor performance of the CNN, additional training data of 1240 samples was added to see if the performance could be improved. However, the additional training data was

not gathered by hand in this study, but rather with ATAS for Publication III. This also presented the possibility to study how comparable the sampling methods are in terms of classification. LDA classification was also conducted on FSFS-selected features to demonstrate the performance in a substantially lower feature space.

3.5 Intraoperative tissue identification

After the studies in laboratory conditions had produced sufficient proof that the concept of tissue identification from surgical smoke works, *in vivo* tissue identification was the next step. In terms of the research setup, the final study was also the most ambitious. The real-time tissue analysis system was tested in a clinical trial with 20 breast cancer patients with approval from the local Ethics Committee of Tampere University Hospital (code R17096). The system was not yet used to guide the surgery in any way or to provide feedback to the operating surgeon. Instead, the aim was to study whether the system could be safely implemented in a normal surgical workflow without delaying or disturbing the operating room staff. Another goal of the research was to study the tissue identification performance with *in vivo* benign tissues. Analysis of malignant tissue was excluded from the study plan, since the inherent goal of cancer operations is to cut around the tumour without making contact. This means that the sample size for malignant tissue would ideally be zero or very low. Thus, there would not be sufficient data for training a classification model.

The measurement system used in the study was similar to the one used in Publication IV. The only hardware differences were the devices used for the smoke evacuation (SafeAir® Smoke Evacuator compact, Stryker Corp, USA) and electrosurgery (Berchtold Elektrotom 530 Electrosurgical Unit, Stryker Corp, USA). On-line analysis was also omitted because it was not possible to create a classification model for the *in vivo* tissues in advance. The adaptive dilution functionality of the electric filter was also omitted as a risk mitigation measure due to the discovery that, in the long term, there were stability issues with the control and, as a result, the dilution did not always function as intended. The system used in the operating room is depicted in Figure 17.

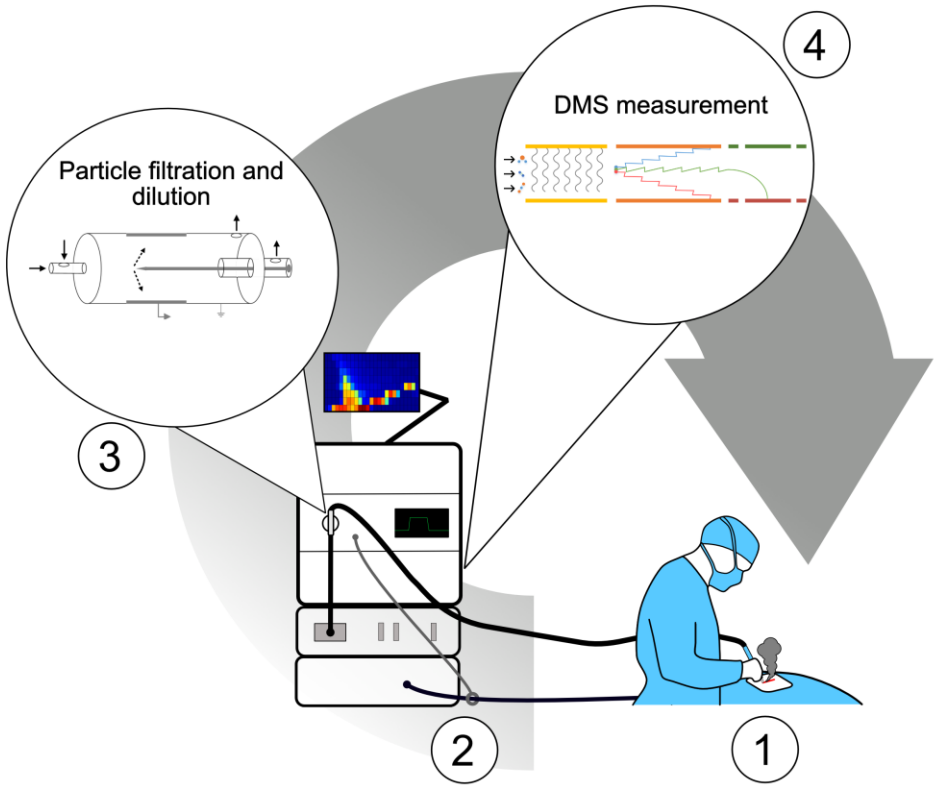


Figure 17. *The final iteration of the real-time tissue analysis system in an operating room environment. 1) Surgical smoke is produced by the diathermy instrument. 2) Induced current from the dispersive electrode triggers the measurement. 3) Surgical smoke sample is taken into the pre-processing unit for particle filtering and dilution. 4) The filtered sample is measured by the DMS sensor, and the result is stored in the internal memory of the system. The duration of one measurement loop is approximately 10 seconds.*

While the system remained largely the same as in the previous study, the rest of the research setup was almost completely changed. The annotation of tissues could no longer rely on careful planning and purposeful incisions on certain tissues. The aim was to minimise the disturbance to the surgeries, which meant that there was no control over the sampling. Instead, the annotation of tissue types for classification required a novel approach, which relied on capturing the full operation with a head-mounted camera (Pupil Core, Pupil Labs GmbH, Germany), worn by the surgeon. The video data from the surgeries was cut into a separate clip for each DMS

measurement and overlaid with the measurement data of the induced diathermy current and DMS spectrum. In total, this resulted in 1131 video clips. Out of these clips, 72 were then viewed individually by three surgeons, who formed their separate opinion on the operated tissue according to an annotation protocol. This approach was used to study the inter-rater agreement of the annotations. The subset size was determined with statistical power calculations. The annotation protocol determined the possible tissue classes as skin, fat, glandular tissue, connective tissue, muscle, blood and empty. After each individual had annotated the subset of clips, the inter-rater agreement was estimated with the Fleiss' kappa, which is commonly used, especially in psychometric studies, to provide a numerical estimate on the degree of agreement [203,204]. Here, the aim was to assess the reliability of the perceived ground-truth for the classification. After the group subset annotation, the remaining video clips were annotated by only one person.

Besides Fleiss' kappa, a new addition to the data analysis of the study was the one-sample t-test analysis of the duration of surgeries to see whether the duration when using the system was statistically different from the institutional average. Additionally, the DMS spectra among different tissues were examined with the Kolmogorov-Smirnov test to see if there are statistically significant differences in the dispersion plots. The DMS resolution was kept the same as in Publication IV (compensation voltage sweep was set to -0.8–5 V with 25 steps and separation voltage sweep to 340–770 V with 8 steps). The classification of the DMS data was performed with LDA. After the exclusion of erroneous or non-annotatable measurement files, the total number of classifiable measurements was 611. The three main reasons for the extensive exclusion of measurements were corruption of the video data in two surgeries, non-annotatable video data and insufficient sampling. The classification models were cross-validated with the leave-one-patient-out method to get a realistic general estimation of the tissue identification performance with *in vivo* benign tissues. In addition, the classification performance for each individual surgery was evaluated with LOOCV LDA to study the variation of the performance between surgeries.

Despite the many differences in the materials and methods in each study presented in this thesis, they create a natural continuum of development and evolution for the tissue identification system. A summary of the general aspects of the research setups for each study are presented in Table 4.

Table 4. *Overview of the research setups of the studies presented in this thesis.*

Study	Material & N	System	DMS resolution	Duration	Data analysis
I	Porcine tissue (10 types) N = 700 (60+600+40)	Autosampler, Particle filtration, Fixed dilution	U_C: -0.8–9 V, 40 steps U_{SV}: 440–700 V, 4 steps Total: 320 values (160×2)	Measurement: 12 s Cleansing: 180 s Total: 192 s	LDA (LOOCV, 10-fold CV, and holdout set)
II	Human breast tissue (*ex vivo*) N = 304	Autosampler, Particle filtration, Fixed dilution, Sample chamber	U_C: -0.8–9 V, 90 steps U_{SV}: 440–700 V, 18 steps Total: 1620 values	Measurement: 60 s Cleansing: 60 s Total: 120 s	LDA (10-fold CV and holdout set) FSFS ROC
III	Porcine tissue, (8 types) N = 3418	Autosampler, Particle filtration, Fixed dilution, Needle electrode	U_C: -0.8–5 V, 25 steps U_{SV}: 340–740 V, 8 steps Total: 200 values	Measurement: 6 s Cleansing: 10–30 s Total: 16–36 s	K-means clustering LDA (10-fold CV) FSFS
IV	Porcine tissue, (2 types) N = 1159 (1059+100)	Freehand, Particle filtration, Adaptive dilution	U_C: -0.8–5 V, 25 steps U_{SV}: 340–740 V, 8 steps Total: 200 values	Measurement: 9 s Cleansing: 9 s Total: 18 s	LDA CNN (10-fold CV and an external test set) FSFS
V	Human breast tissue (*in vivo*) N = 611	Freehand, Particle filtration, Fixed dilution	U_C: -0.8–5 V, 25 steps U_{SV}: 340–740 V, 8 steps Total: 200 values	Measurement: 10 s Cleansing: varying Total: varying	LDA (LOOCV and leave-one-patient-out) One sample t-test Kolmogorov-Smirnov Fleiss' Kappa

4 RESULTS AND DISCUSSION

The results in each study provided insight into the performance of the tissue analysis system and highlighted potential areas of improvement. In this chapter, the main results and their key implications are presented. In addition, the limitations of each study are discussed.

4.1 Porcine tissue studies

The results of the porcine tissue studies illustrated the potential of the system in terms of tissue identification. The system used in the first study was unrefined in terms of tubing lengths, sampling and the overall placement of the parts. This resulted in the need to implement a substantial waiting period of approximately 3 minutes to cleanse the system before a new measurement could be made. Still, in the first study, the classification accuracies for porcine tissues were promising. For the initial test set of ten samples from five tissues measured in one sitting, the LOOCV LDA classification accuracy was 93%. When the number of classes was expanded to ten in the second phase, the classification accuracy was 95% with 10-fold cross-validation and 93% with the 70–30 holdout method. In phase III with an additional ten samples from four tissue types, the classification accuracy was 95%. For a proof-of-concept study, these numbers seemed exceptionally good. However, it is worth noting that the study design and analysis in Publication I were suboptimal in terms of the generalisability of the results. The results from the classification of phase II in Publication I are especially problematic, since sample randomisation was not considered, thus resulting in an overly optimistic interpretation. This was discussed in Publication III and partly corrected in the study design. Still, LOOCV and 10-fold cross-validation were a poor choice of methodology from the author's part, which must be addressed. As discussed in the theoretical background section, by selecting the cross-validation partitions randomly from within measurement sets that are not completely independent, the possible variations in the measurement conditions and their effect on the measurement results are included in both the training and the test sets, which means that the results are biased and do not provide an accurate estimation of the performance of the system for fully independent new data. The descriptions of the

methodology of Publications I, II and III provide information on the utilised cross-validation method but do not properly address the issues of generalisability. In Publications IV and V, the data analysis and classification were better executed in this regard.

In Publication III, the results for the classification of the tissues that had macroscopically distinct tissue types enabled the imaging of the tissues with adequate accuracy using the posterior probability scores. Thus, the concept of a pathological imaging was demonstrated. Measurements made on flank pieces were most successfully classified as skeletal muscle and intramuscular fat, with an accuracy of 91.8%. Grey and white matter in brain samples were also identified with over 90% accuracy. The worst classification performance was observed with renal cortex and renal pelvis tissue, with an accuracy of 70.9%. Example illustrations of tissue imaging based on the classification scores are presented in Figure 18.

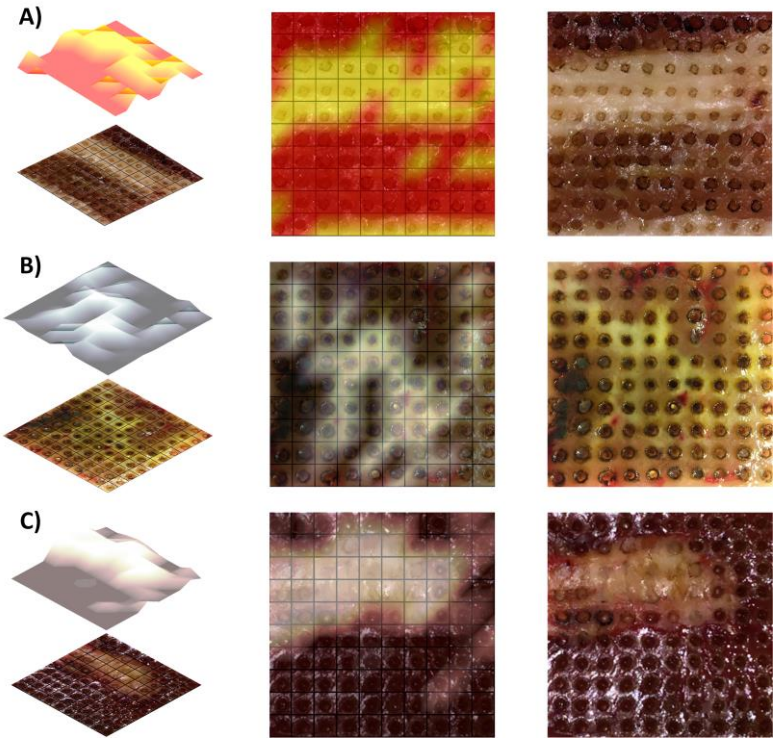


Figure 18. *Tissue imaging. A) Porcine flank sample. B) Porcine brain sample. C) Porcine kidney sample. Reprinted with permission from Elsevier: Elsevier, Experimental and Molecular Pathology, Differential mobility spectrometry imaging for pathological applications, Kontunen et al., Copyright © 2020 Elsevier Inc. All rights reserved.*

The tissue imaging with K-means clustering was not as accurate as with LDA, but even with the unsupervised clustering, the created images resembled the actual measured tissue, especially in the flank pieces. Overall, the misclassifications in the binary setting can be partly due to the depth of the sampling. Some of the 3-millimetre-deep automated cuts may have inadvertently vaporised both tissues, thus creating a mixed signal. Also, one likely reason for the poor performance in the kidney classification was the difficulty in sampling the renal pelvis. Some of the cuts made onto renal pelvis tissue vaporised the tissue effectively, but some cuts only made the needle electrode stick to the sample without effective vaporisation. This likely created high within-class variation in the DMS responses. Furthermore, when considering an actual pathological application, the achievable spatial resolution of 1 mm with the needle electrode is not enough. To image microscopic tumours, µm-scale spatial resolution is needed. Another issue with the presented method was its destructiveness. In order for the technology to be used for pathological analysis as is, duplicate tissue slice samples from the same area would be required. One would be used for the ATAS analysis and the other for more accurate histopathological analysis. This would obviously decrease the usability of the method. However, all of these problems related to the sampling could be overcome with minimally destructive laser ablation similar to the MS-based cancer margin assessment methods. With laser sampling, the spatial resolution, tissue damage, and smoke contamination could be minimised, and it is likely that the classification results would also improve.

The 8-class classification results in Publication III revealed that the results presented in Publication I were likely overly optimistic. The overall accuracy was 81%. The FSFS analysis revealed that distinct features in the DMS spectra can be used to identify the different tissue types. On average, the number of features selected to identify the tissues was 8. This means that the majority of the spectrum features are, in fact, redundant in terms of information content and that if the DMS device itself would not require the systematic measurement of all U_C-U_{SV} pairs between certain limit values, the DMS measurement could be made significantly faster, without compromising the identification performance. This same result could be inferred from the FSFS results in Publication IV, where the classification with the FSFS-selected features were in fact slightly better (93.2%) than with the full spectra (93.1%).

In Publication IV, the problem was simplified to a binary classification of skeletal muscle and adipose tissue, since the focus of the study was more on the overall feasibility of real-time tissue analysis. The classification results with both LDA

(93.1%) and CNN (93.2%) were good in the 10-fold cross-validated initial set of 1059 samples. However, the CNN classification performance dropped to guess level (50%), when the independent validation set was used as the test set. This implies that the CNN model was heavily overfitted to the features found in the initial dataset. The most probable reason as to why the classification failed completely with CNN was the disassembly and reassembly of the DMS sensor core due to the physical blockage in the pneumatics. It is likely that the tightness of the screws or electrode placements were altered ever so slightly after the reassembly, which could have had a small impact on the produced dispersion plots that was amplified by the overfitting CNN model. As a simpler linear model, the LDA model trained with the initial set was more robust as regards the potential change, and when it was used as the real-time classification model during the measurement of the validation set, the achieved classification accuracy was 87%. Another interesting aspect of the results in Publication IV were the classification performances with the models trained with additional data that were not comparable in terms of the sampling method. When the ATAS data for muscle and fat from Publication III were added for model training, the CNN's classification accuracy of the validation set was 88% and the LDA achieved a classification accuracy of 96%. These results illustrate the importance of having training samples that are measured in as varied circumstances as possible. This way, the features that generalise to the classification problem are truly indicative of the class differences rather than the environmental conditions. All in all, the results in Publication IV demonstrate, albeit in a simplified setting, that intraoperative tissue identification is possible with the proposed system.

Even though some of the presented results suffered from poor generalisability due to the lack of external validation datasets, the aim of identifying porcine tissues based on the DMS measurement of surgical smoke was achieved. The porcine tissues used in the studies were frozen in a normal freezer prior to analysis, which also potentially decreases the generalisability of the results, since slow freezing has been shown to cause myofibrillar damage and decrease the water holding capacity of porcine tissue [205]. Still, the tissues were typically not stored for more than a couple of weeks before measurements, so it is likely that extensive lipid oxidation did not take place and that the molecular content of the tissues remained comparable to fresh tissues. In earlier MS studies regarding animal tissue identification from surgical smoke, the molecular differences have been attributed to various lipid species and their degradation products [194,206]. However, lipids only account for approximately 5% of the dry mass of mammalian cells [44]. The relative abundance of proteins, carbohydrates and water also vary among tissue types and contribute to the molecular

structure [25,207]. The DMS spectrum of the vaporised tissue entails these differences and forms a view of the mixture of ion species that are present in the sample. The analytical performance of DMS is at a lower level compared to MS, but the studies presented in this thesis indicate that accurate differentiation of the tissue types is possible. Table 5 provides a summary of the tissue identification results, their key implications and the main limitations of the porcine tissue studies.

Table 5. *Tissue identification results for the porcine tissue studies.*

Study	Classification accuracy	Key finding or implication	Main limitations
I	**Phase I**: 93% **Phase II**: 95% **Phase III**: 95%	Operation principle is feasible	Rudimentary system, lack of sample randomisation (especially in phase II), no independent validation dataset
III	**Binary classification of heterogeneous tissues**: 71%–92% **Full set**: 81%	Imaging of heterogeneous tissues is feasible	Low spatial resolution, duration, destructiveness, no independent validation dataset
IV	**Initial set with cross-validation**: LDA: 93%; CNN: 93% **With FSFS**: LDA: 93% **Independent test set**: LDA: 87%; CNN: 50% **With additional training**: LDA: 96%; CNN: 88%	Real-time tissue identification is feasible	Simplified classification setup, ENVI-AMC

4.2 Human breast cancer studies

The first study with fresh breast cancer samples, presented in Publication II, was successful in demonstrating the feasibility of clinically relevant tissue identification with the DMS-based system. The 10-fold cross-validated LDA model reached a mean classification accuracy of 86.5%, with a sensitivity of 80.1% and specificity of 89.9% for the binary classification of malignant and benign tissues. The malignant tissues were also classified based on cancer type into ductal, lobular and invasive micropapillary carcinoma with an accuracy of 80.2%. In addition, FSFS was performed for the DMS data. As in the porcine tissue studies, the feature selection again demonstrated that a similar level of classification performance (82.4%) can be achieved by using only a fraction of the features in the dispersion plot (7 out of 1620). This again highlighted the need for a DMS system that would allow the measurement of only selected pairs of U_C and U_{SV} voltages. Especially in this study, where the measurement of the spectra lasted for approximately one minute, the time saved by measuring only the FSFS-selected features would have been substantial. The ROC curve plotted of the classification further demonstrated the performance. It is worth noting that the confidence intervals are quite wide, meaning that the performance has some variation depending on the cross-validation partitions. Also, the lack of a completely independent validation set means that the results are not guaranteed to fully generalise to new data. In addition, the system used in the study was still somewhat unreliable, which caused erroneous measurements that needed to be discarded. Still, as the first study in breast tissue identification, the results were promising.

In this respect, the classification results from the final study were disappointing. While there were statistically significant differences between the spectra of the tissue types, the classification accuracy with leave-one-patient-out cross-validated LDA model was only 44% for four tissue types. Also, the LOOCV LDA for each surgery separately revealed significant variation in the classification accuracy, with a range of 37%–100%. The most likely reasons for the poor classification performance were the delays in the DMS measurement and challenges in the annotation of the tissues. In some cases, the focus of the video was off, or the operation area was obscured by hands, blood or the bright operating room lighting. Annotation based on video data is of course much more unreliable than histopathological analysis, or even visual examination in a controlled laboratory environment. Therefore, the samples for the classes that the machine learning models consider to be absolute ground truths can be mislabelled. To study the annotation process, the inter-rater agreement was

estimated for three annotators using the Fleiss' kappa metric. Based on the metric, the agreement could be interpreted as "Moderate", which infers that, while the annotators agreed on some labels, much of the process was prone to subjective interpretation. The delay in the start of the DMS measurement likely caused some degree of uncertainty in the assessment of the annotators. Therefore, in this study, the decision to use ENVI-AMC as the DMS device was questionable. The device was a prototype initially designed for clean room monitoring, and it was not designed for rapid measurements. This became apparent upon discovering a software bug that caused part of the measurement data to be corrupted if the duration of the measurement was shortened too much. To ensure that all measurement data were valid, the delay between the trigger signal and the measurement could not be removed. In the other studies, the limitations of the DMS system were not as crucial, and, in retrospect, the system was in fact very reliable throughout all of these measurements.

Still, in hindsight, it can be argued that the operating room trial was conducted too early in the development of the system and that it would have been better to wait until a more suitable sensor was available. Also, the results would likely have been better, if the surgeon had been instructed to purposefully operate certain types of tissues at some points, but this would have gone against the aim of minimising the disturbance to the normal surgical workflow. Another addition that could have improved the classification results is environmental control. The temperature and humidity between the surgeries varied, which likely influenced the measured outputs. However, despite the poor classification results, the study was an important step forward, since it highlighted the areas of improvement that were still needed. Furthermore, the study provided positive results in terms of the applicability of the system to a normal surgical workflow. There was no effect on the complication rate or the duration of the surgeries when the system was used. All in all, the human sample studies were not as successful as the porcine tissue studies in demonstrating the capability of the system in tissue identification. Table 6 provides a summary of the results from the human tissue measurements.

Table 6. *Tissue identification results for the breast cancer studies.*

Study	Main results	Key finding or implication	Main limitations
II	**Malignant vs. benign**: classification accuracy = 87%, sens = 80%, spec = 90% **With FSFS**: classification accuracy = 82% **AUC**: 0.90 **Cancer type classification**: classification accuracy = 80%	Technology is clinically feasible in terms of *ex vivo* tissue identification performance	Moderately frequent system malfunctions, no independent validation dataset
V	**LDA classification accuracy**, Leave-one-patient out: 44% LOOCV: 37%–100% **Inter-rater agreement in annotation**: Moderate **Causes complications**: No **Causes delays**: No	Technology is not ready for *in vivo* use, but the system is applicable to the surgical workflow	Uncertainty in reference labelling, ENVI-AMC, delays, environmental variation was not compensated or controlled

The molecular composition of surgical smoke between cancerous and benign tissue has been shown to differ in MS studies, especially in terms of phospholipids and triglycerides [107,108,119]. The underlying premise of the differences in metabolism between benign and malignant tissue, including the Warburg effect that generalises to almost every type of cancer, can at least partly explain the root cause of the differences in the molecular profile between cancerous and benign tissue. With DMS analysis, the determination of certain metabolites from a complex matrix is not feasible, which creates a degree of uncertainty for the identification of the samples. In the future, it will be vital to test the tissue identification simultaneously with the DMS-based system and a scientifically validated reference method, such as REIMS. Only then can conclusive evidence be provided that the DMS-based method can compete in the field of (intraoperative) tissue identification. Currently, the DMS-based technology is in a state where the preliminary results have been promising but the maturity of the technology is not yet at a level that enables its utilisation in clinical practice. The next major steps forward include the accurate compensation of the environmental conditions that would enable stable sample humidity and a DMS sensor that could measure selected U_C-U_{SV} pairs rapidly.

5 CONCLUSIONS

This thesis presented five main research questions. These questions were answered through the studies presented in the thesis, and in five publications. The final compiled conclusions to the presented research questions are as follows:

I. Is DMS analysis of surgical smoke a feasible method for *ex vivo* tissue identification in a porcine model?

- **Yes**. The classification accuracies achieved with porcine tissues in the first study show that DMS-based analysis of tissue smoke can be used to identify normal porcine tissue types in laboratory conditions. However, the limitations of the study need to be acknowledged, and the numerical results presented originally in Publication I can be viewed as overly optimistic.

II. Can DMS-based *ex vivo* tissue identification be utilised with clinically relevant human breast tissue?

- **Yes**. The diagnostic performance measures achieved in the second study with histologically verified benign and malignant human breast samples demonstrate that DMS-based analysis of tissue smoke can be used to identify breast malignancies in laboratory conditions. The reported results in Publication II can be partially overly optimistic due to a lack of independent validation dataset, but, overall, the performance shows promise towards a clinical application.

III. Is DMS-based *ex vivo* tissue identification feasible as an imaging method for pathological applications?

- **Yes**. The DMS measurement of surgical smoke and the subsequent classification scores enable the imaging of heterogeneous tissues. With the further improvements to the system presented in Publication III regarding spatial resolution and duration of analysis, and with validation with clinically relevant tissues, the technology could be implemented as a complimentary method to guide histopathological analysis.

IV. Can DMS-based tissue identification be performed rapidly enough to enable intraoperative use?

- **Yes**. The measurements and classification results in Publication IV demonstrate that tissue identification, at least in a simplified laboratory setting, can be done within a timeframe that allows intraoperative use for tissue identification. However, for improved applicability to the surgical workflow, the delays in measurement and analysis need to be decreased further.

V. Is DMS-based tissue identification feasible in a surgical setting?

- **Not yet**. The results achieved in Publication V were not comparable to previous results obtained in laboratory conditions in terms of diagnostic accuracy, and the system would not be reliable enough to aid in margin assessment. The difficulties in intraoperative tissue annotation and the long delays between subsequent measurements due to suboptimal sensor hardware are a likely explanation for the poor results. Further technological development is required to fix the limitations regarding sensor hardware, environmental control and the reliability of the annotation in the study, after which the research question should be revisited to provide a conclusive answer.

Overall, despite the inconclusive answer to the last research question and some over-optimistic result representation due to the lack of external validation datasets in the published work, the research presented in this thesis was successful in providing positive answers to the original research questions. The development of a DMS-based tissue identification system saw significant leaps forward towards potential commercialisation. However, the technology is not yet mature enough to be implemented in clinical practice. In particular, the research demonstrated the need for a rapid DMS sensor that would be capable of measuring selected voltage pairs. Even with a new sensor, the roadmap towards a fully realised medical device has many technological obstacles to overcome, but the research in this thesis demonstrated that DMS analysis of surgical smoke has true potential to be developed into a helpful tool in cancer margin assessment and pathological applications. In its ideal future form, the technology presented in this thesis could save millions of people from avoidable cancer reoperations and billions of dollars in healthcare costs.

BIBLIOGRAPHY

1. American Cancer Society. Basic Cancer Facts. In: Global Cancer Facts & Figures 4th Edition. Atlanta: American Cancer Society; 2018. p. 1–11.

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021 Feb 4;71(3):209–49.

3. American Cancer Society. Breast Cancer Treatment. In: Breast Cancer Facts & Figures 2019-2020. Atlanta: American Cancer Society; 2019. p. 23–6.

4. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2019 Aug;30(8):1194–220.

5. Heil J, Czink E, Golatta M, Schott S, Hof H, Jenetzky E, et al. Change of aesthetic and functional outcome over time and their relationship to quality of life after breast conserving therapy. Eur J Surg Oncol. 2011 Feb;37(2):116–21.

6. Waljee JF, Hu ES, Ubel PA, Smith DM, Newman LA, Alderman AK. Effect of esthetic outcome after breast-conserving surgery on psychosocial functioning and quality of life. J Clin Oncol. 2008 Jul;26(20):3331–7.

7. Volders JH, Negenborn VL, Haloua MH, Krekel NMA, Jóźwiak K, Meijer S, et al. Cosmetic outcome and quality of life are inextricably linked in breast-conserving therapy. J Surg Oncol. 2017 Jun;115(8):941–8.

8. Cochrane RA, Valasiadou P, Wilson ARM, Al-Ghazal SK, Macmillan RD. Cosmesis and satisfaction after breast-conserving surgery correlates with the percentage of breast volume excised. Br J Surg. 2003 Dec;90(12):1505–9.

9. Moran MS, Schnitt SJ, Giuliano AE, Harris JR, Khan SA, Horton J, et al. Society of surgical oncology-American society for radiation oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in stages i and II invasive breast cancer. Int J Radiat Oncol Biol Phys. 2014 Mar;88(3):553–64.

10. Niinikoski L, Leidenius MHK, Vaara P, Voynov A, Heikkilä P, Mattson J, et al. Resection margins and local recurrences in breast cancer: Comparison

between conventional and oncoplastic breast conserving surgery. Eur J Surg Oncol. 2019 Jun;45(6):976–82.

11. Landercasper J, Borgert AJ, Fayanju OM, Cody H, Feldman S, Greenberg C, et al. Factors Associated with Reoperation in Breast-Conserving Surgery for Cancer: A Prospective Study of American Society of Breast Surgeon Members. Ann Surg Oncol. 2019 Jul;26(10):3321–36.

12. Kouzminova NB, Aggarwal S, Aggarwal A, Allo MD, Lin AY. Impact of initial surgical margins and residual cancer upon re-excision on outcome of patients with localized breast cancer. Am J Surg. 2009 Dec;198(6):771–80.

13. Grant Y, Al-Khudairi R, St John E, Barschkett M, Cunningham D, Al-Mufti R, et al. Patient-level costs in margin re-excision for breast-conserving surgery. Br J Surg. 2019 Mar 1;106(4):384–94.

14. Virtanen J, Hokkinen L, Karjalainen M, Kontunen A, Vuento R, Numminen J, et al. In vitro detection of common rhinosinusitis bacteria by the eNose utilising differential mobility spectrometry. Eur Arch Oto-Rhino-Laryngology. 2018;275(9):2273–9.

15. Niemi RJ, Roine AN, Eräviita E, Kumpulainen PS, Mäenpää JU, Oksala N. FAIMS analysis of urine gaseous headspace is capable of differentiating ovarian cancer. Gynecol Oncol. 2018 Dec 1;151(3):519–24.

16. Alberts B, Johnson A, Raff M, Lewis J, Roberts K, Walter P, et al. The Maintenance of DNA Sequences. In: Molecular Biology of the Cell. 5th ed. New York: Garland Science; 2007. p. 263–5.

17. Kim N, Jinks-Robertson S. Transcription as a source of genome instability. Nat Rev Genet. 2012;13(3):204–14.

18. Cristian T, Bert V. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science. 2015 Jan 2;347(6217):78–81.

19. Alberts B, Bray D, Hopkin K, Johnson A, Lewis J, Raff M, et al. Cancer. In: Essential Cell Biology. 3rd ed. New York, NY, USA: Garland Science; 2010. p. 717–31.

20. Wong RSY. Apoptosis in cancer: from pathogenesis to treatment. J Exp Clin Cancer Res. 2011 Sep 26;30(87).

21. Luffarelli P, Manna E, Fortunato L. Epidemiology and Risk Factors BT - Ductal Carcinoma in Situ of the Breast. In: Mariotti C, editor. Cham: Springer International Publishing; 2018. p. 23–37.

22. Pluchinotta AM. History of the DCIS of the Breast and the Evolution of Knowledge Based on Ductal Tree Anatomy BT - Ductal Carcinoma in Situ of the Breast. In: Mariotti C, editor. Cham: Springer International

Publishing; 2018. p. 1–22.

23. Mariotti C, Raffaeli E. The Surgical Treatment of DCIS: from Local Excision to Conservative Breast Surgery and Conservative Mastectomies BT - Ductal Carcinoma in Situ of the Breast. In: Mariotti C, editor. Cham: Springer International Publishing; 2018. p. 107–42.

24. Cooper GM. The Molecules of Cells. In: The Cell: a Molecular Approach. 8th ed. New York, NY, USA: Oxford University Press; 2019. p. 45–62.

25. Woodard HQ, White DR. The composition of body tissues. Br J Radiol. 1986;59(708):1209–18.

26. Jögi A, Vaapil M, Johansson M, Påhlman S. Cancer cell differentiation heterogeneity and aggressive behavior in solid tumors. Ups J Med Sci. 2012 Feb 29;117(2).

27. Ilyin SE, Belkowski SM, Plata-Salamán CR. Biomarker discovery and validation: technologies and integrative approaches. Trends Biotechnol. 2004;22(8):411–6.

28. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips K-A, Mooij TM, Roos-Blom M-J, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. JAMA. 2017 Jun;317(23):2402–16.

29. Moreau K, Dizin E, Ray H, Luquain C, Lefai E, Foufelle F, et al. BRCA1 Affects Lipid Synthesis through Its Interaction with Acetyl-CoA Carboxylase *. J Biol Chem. 2006 Feb;281(6):3172–81.

30. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001 Mar 1;69(3):89–95.

31. Füzéry AK, Levin J, Chan MM, Chan DW. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. Clin Proteomics. 2013;10(1):13.

32. Chen F, Chandrashekar DS, Varambally S, Creighton CJ. Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers. Nat Commun. 2019;10(1):5679.

33. Ludwig JA, Weinstein JN. Biomarkers in Cancer Staging, Prognosis and Treatment Selection. Nat Rev Cancer. 2005;5(11):845–56.

34. Warburg O. The Metabolism of Carcinoma Cells. J Cancer Res. 1925 Mar 1;9(1):148–63.

35. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science. 2009 May 22;324(5930):1029–33.

36. Koppenol WH, Bounds PL, Dang C V. Otto Warburg's contributions to current concepts of cancer metabolism. Nat Rev Cancer. 2011;11(5):325–37.

37. DeBerardinis RJ, Chandel NS. We need to talk about the Warburg effect. Nat Metab. 2020;2(2):127–9.

38. Newsholme EA, Crabtree B, Ardawi MSM. The role of high rates of glycolysis and glutamine utilization in rapidly dividing cells. Biosci Rep. 1985 May 1;5(5):393–400.

39. Alfarouk KO, Schwartz CTS and L. The Warburg Effect and the Hallmarks of Cancer. Vol. 17, Anti-Cancer Agents in Medicinal Chemistry. 2017. p. 164–70.

40. Chen P-H, Cai L, Huffman K, Yang C, Kim J, Faubert B, et al. Metabolic Diversity in Human Non-Small Cell Lung Cancer Cells. Mol Cell. 2019;76(5):838–51.

41. Lim JY, Kwan HY. Roles of Lipids in Cancer. In: Baez RV, editor. Advances in Lipid Metabolism. 1st ed. London: IntechOpen; 2020. p. 1–20.

42. Long J, Zhang C-J, Zhu N, Du K, Yin Y-F, Tan X, et al. Lipid metabolism and carcinogenesis, cancer development. Am J Cancer Res. 2018 May 1;8(5):778–91.

43. Szlasa W, Zendran I, Zalesińska A, Tarek M, Kulbacka J. Lipid composition of the cancer cell membrane. J Bioenerg Biomembr. 2020;52(5):321–42.

44. Alberts B, Johnson A, Raff M, Lewis J, Roberts K, Walter P, et al. The Chemical Components of a Cell. In: Molecular Biology of the Cell. 5th ed. New York: Garland Science; 2007. p. 45–65.

45. Alberts B, Johnson A, Raff M, Lewis J, Roberts K, Walter P, et al. The Lipid Bilayer. In: Molecular Biology of the Cell. 5th ed. New York: Garland Science; 2007. p. 617–29.

46. Hilvo M, Denkert C, Lehtinen L, Müller B, Brockmöller S, Seppänen-Laakso T, et al. Novel Theranostic Opportunities Offered by Characterization of Altered Membrane Lipid Metabolism in Breast Cancer Progression. Cancer Res. 2011 May 1;71(9):3236–45.

47. Dzendrowskyj TE, Noyszewski EA, Beers J, Bolinger L. Lipid composition changes in normal breast throughout the menstrual cycle. Magn Reson Mater Physics, Biol Med. 1997;5(2):105–10.

48. Tan J, Le A. The Heterogeneity of Breast Cancer Metabolism BT - The Heterogeneity of Cancer Metabolism. In: Le A, editor. Cham: Springer International Publishing; 2021. p. 89–101.

49. Azordegan N, Fraser V, Le K, Hillyer LM, Ma DWL, Fischer G, et al. Carcinogenesis alters fatty acid profile in breast tissue. Mol Cell Biochem. 2013;374(1):223–32.

50. Vaysse P-M, Kooreman LFS, Engelen SME, Kremer B, Olde Damink SWM, Heeren RMA, et al. Stromal vapors for real-time molecular guidance of breast-conserving surgery. Sci Rep. 2020;10(20109).

51. Sladden MJ, Nieweg OE, Howle J, Coventry BJ, Thompson JF. Updated evidence-based clinical practice guidelines for the diagnosis and management of melanoma: definitive excision margins for primary cutaneous melanoma. Med J Aust. 2018 Feb 1;208(3):137–42.

52. Nahhas AF, Scarbrough CA, Trotter S. A Review of the Global Guidelines on Surgical Margins for Nonmelanoma Skin Cancers. J Clin Aesthet Dermatol. 2017 Apr;10(4):37–46.

53. Schmoll HJ, Van Cutsem E, Stein A, Valentini V, Glimelius B, Haustermans K, et al. ESMO Consensus Guidelines for management of patients with colon and rectal cancer. A personalized approach to clinical decision making. Ann Oncol. 2012 Oct 1;23(10):2479–516.

54. Orosco RK, Tapia VJ, Califano JA, Clary B, Cohen EEW, Kane C, et al. Positive Surgical Margins in the 10 Most Common Solid Cancers. Sci Rep. 2018 Apr 9;8(5686).

55. Wittekind C, Compton C, Quirke P, Nagtegaal I, Merkel S, Hermanek P, et al. A uniform residual tumor (R) classification. Cancer. 2009 Aug 1;115(15):3483–8.

56. Alexander BM, Cloughesy TF. Adult Glioblastoma. J Clin Oncol. 2017 Jun 22;35(21):2402–9.

57. Ecclestone BR, Bell K, Abbasi S, Dinakaran D, van Landeghem FKH, Mackey JR, et al. Improving maximal safe brain tumor resection with photoacoustic remote sensing microscopy. Sci Rep. 2020;10(17211).

58. American Cancer Society. Selected Cancers. In: Global Cancer Facts & Figures 4th Edition. American Cancer Society; 2018. p. 12–47.

59. Darby S, McGale P, Correa C, Taylor C, Arriagada R, Clarke M, et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: Meta-analysis of individual patient data for 10 801 women in 17 randomised trials. Lancet. 2011;378(9804):1707–16.

60. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER, et al. Twenty-Year Follow-up of a Randomized Trial Comparing Total Mastectomy, Lumpectomy, and Lumpectomy plus Irradiation for the Treatment of Invasive Breast Cancer. N Engl J Med. 2002 Oct;347(16):1233–41.

61. Agarwal S, Pappas L, Neumayer L, Kokeny K, Agarwal J. Effect of breast conservation therapy vs mastectomy on disease-specific survival for early-stage breast cancer. JAMA Surg. 2014 Mar;149(3):267–74.

62. Lagendijk M, van Maaren MC, Saadatmand S, Strobbe LJA, Poortmans PMP, Koppert LB, et al. Breast conserving therapy and mastectomy revisited: Breast cancer-specific survival and the influence of prognostic factors in 129,692 patients. Int J Cancer. 2018 Jan;142(1):165–75.

63. Gradishar WJ, Anderson BO, Abraham J, Aft R, Agnese D, Allison KH, et al. Breast Cancer, Version 3.2020, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Cancer Netw. 2020;18(4):452–78.

64. Morrow M, Van Zee KJ, Solin LJ, Houssami N, Chavez-MacGregor M, Harris JR, et al. Society of Surgical Oncology–American Society for Radiation Oncology–American Society of Clinical Oncology Consensus Guideline on Margins for Breast-Conserving Surgery With Whole-Breast Irradiation in Ductal Carcinoma In Situ. J Clin Oncol. 2016 Aug 15;34(33):4040–6.

65. Krekel N, Zonderhuis B, Muller S, Bril H, Van Slooten HJ, De Lange De Klerk E, et al. Excessive resections in breast-conserving surgery: A retrospective multicentre study. Breast J. 2011 Nov;17(6):602–9.

66. Biganzoli L, Marotti L, Hart CD, Cataliotti L, Cutuli B, Kühn T, et al. Quality indicators in breast cancer care: An update from the EUSOMA working group. Eur J Cancer. 2017 Nov;86:59–81.

67. Landercasper J, Attai D, Atisha D, Beitsch P, Bosserman L, Boughey J, et al. Toolbox to Reduce Lumpectomy Reoperations and Improve Cosmetic Outcome in Breast Cancer Patients: The American Society of Breast Surgeons Consensus Conference. Ann Surg Oncol. 2015;22:3174–83.

68. McCahill LE, Single RM, Aiello Bowles EJ, Feigelson HS, James TA, Barney T, et al. Variability in Reexcision Following Breast Conservation Surgery. JAMA. 2012 Feb 1;307(5):467–75.

69. Wilke LG, Czechura T, Wang C, Lapin B, Liederbach E, Winchester DP, et al. Repeat Surgery After Breast Conservation for the Treatment of Stage 0 to II Breast Carcinoma: A Report From the National Cancer Data Base, 2004-2010. JAMA Surg. 2014 Dec 1;149(12):1296–305.

70. Jeevan R, Cromwell DA, Trivella M, Lawrence G, Kearins O, Pereira J, et al. Reoperation rates after breast conserving surgery for breast cancer among women in England: Retrospective study of hospital episode statistics. BMJ. 2012;345(7869):1–9.

71. Landercasper J, Whitacre E, Degnim AC, Al-Hamadani M. Reasons for Re-Excision After Lumpectomy for Breast Cancer: Insight from the American Society of Breast Surgeons MasterySM Database. Ann Surg Oncol. 2014;21(10):3185–91.

72. Pataky RE, Baliski CR. Reoperation costs in attempted breast-conserving surgery: A decision analysis. Curr Oncol. 2016;23(5):314–21.

73. Nunez A, Jones V, Schulz-Costello K, Schmolze D. Accuracy of gross intraoperative margin assessment for breast cancer: experience since the SSO-ASTRO margin consensus guidelines. Sci Rep. 2020;10(17344).

74. WILSON LB. A METHOD FOR THE RAPID PREPARATION OF FRESH TISSUES FOR THE MICROSCOPE. J Am Med Assoc. 1905 Dec 2;XLV(23):1737.

75. Gal AA, Cagle PT. The 100-Year Anniversary of the Description of the Frozen Section Procedure. JAMA. 2005 Dec 28;294(24):3135–7.

76. St John ER, Al-Khudairi R, Ashrafian H, Athanasiou T, Takats Z, Hadjiminas DJ, et al. Diagnostic accuracy of intraoperative techniques for margin assessment in breast cancer surgery a meta-analysis. Ann Surg. 2017;265(2):300–10.

77. Jorns JM, Visscher D, Sabel M, Breslin T, Healy P, Daignaut S, et al. Intraoperative frozen section analysis of margins in breast conserving surgery significantly decreases reoperative rates: One-year experience at an ambulatory surgical center. Am J Clin Pathol. 2012;138(5):657–69.

78. Garcia MT, Mota BS, Cardoso N, Martimbianco ALC, Ricci MD, Carvalho FM, et al. Accuracy of frozen section in intraoperative margin assessment for breast-conserving surgery: A systematic review and meta-analysis. PLoS One. 2021 Mar 18;16(3).

79. Esbona K, Li Z, Wilke LG. Intraoperative Imprint Cytology and Frozen Section Pathology for Margin Assessment in Breast Conservation Surgery: A Systematic Review. Ann Surg Oncol. 2012;19(10):3236–45.

80. Singletary SE. Surgical margins in patients with early-stage breast cancer treated with breast conservation therapy. Am J Surg. 2002;184(5):383–93.

81. Bakhshandeh M, Tutuncuoglu SO, Fischer G, Masood S. Use of imprint cytology for assessment of surgical margins in lumpectomy specimens of

breast cancer patients. Diagn Cytopathol. 2007 Oct 1;35(10):656–9.

82. Weinberg E, Cox C, Dupont E, White L, Ebert M, Greenberg H, et al. Local recurrence in lumpectomy patients after imprint cytology margin evaluation. Am J Surg. 2004;188(4):349–54.

83. Laucirica R. Intraoperative Assessment of the Breast: Guidelines and Potential Pitfalls. Arch Pathol Lab Med. 2005 Dec 1;129(12):1565–74.

84. Tamanuki T, Namura M, Aoyagi T, Shimizu S, Suwa T, Matsuzaki H. Effect of Intraoperative Imprint Cytology Followed by Frozen Section on Margin Assessment in Breast-Conserving Surgery. Ann Surg Oncol. 2021;28(3):1338–46.

85. Kaufman CS, Jacobson L, Bachman BA, Kaufman LB, Mahon C, Gambrell L-J, et al. Intraoperative Digital Specimen Mammography: Rapid, Accurate Results Expedite Surgery. Ann Surg Oncol. 2007;14(4):1478–85.

86. Kim SHH, Cornacchi SD, Heller B, Farrokhyar F, Babra M, Lovrics PJ. An evaluation of intraoperative digital specimen mammography versus conventional specimen radiography for the excision of nonpalpable breast lesions. Am J Surg. 2013;205(6):703–10.

87. Chagpar AB, Butler M, Killelea BK, Horowitz NR, Stavris K, Lannin DR. Does three-dimensional intraoperative specimen imaging reduce the need for re-excision in breast cancer patients? A prospective cohort study. Am J Surg. 2015;210(5):886–90.

88. Colakovic N, Zdravkovic D, Skuric Z, Mrda D, Gacic J, Ivanovic N. Intraoperative ultrasound in breast cancer surgery—from localization of non-palpable tumors to objectively measurable excision. World J Surg Oncol. 2018;16(1):184.

89. Moschetta M, Telegrafo M, Introna T, Coi L, Rella L, Ranieri V, et al. Role of specimen US for predicting resection margin status in breast conserving therapy. G Chir. 2015;36(5):201–4.

90. Dashevsky BZ, D'Alfonso T, Sutton EJ, Giambrone A, Aronowitz E, Morris EA, et al. The Potential of High Resolution Magnetic Resonance Microscopy in the Pathologic Analysis of Resected Breast and Lymph Tissue. Sci Rep. 2015;5(17435).

91. Hall NC, Povoski SP, Murrey DA, Knopp M V, Martin EW. Combined approach of perioperative 18F-FDG PET/CT imaging and intraoperative 18F-FDG handheld gamma probe detection for tumor localization and verification of complete tumor resection in breast cancer. World J Surg Oncol. 2007;5(1):143.

92. Balasundaram G, Krafft C, Zhang R, Dev K, Bi R, Moothanchery M, et al. Biophotonic technologies for assessment of breast tumor surgical margins—A review. J Biophotonics. 2021 Jan 1;14(1).

93. Li R, Wang P, Lan L, Lloyd FP, Goergen CJ, Chen S, et al. Assessing breast tumor margin by multispectral photoacoustic tomography. Biomed Opt Express. 2015;6(4):1273–81.

94. Keller MD, Majumder SK, Kelley MC, Meszoely IM, Boulos FI, Olivares GM, et al. Autofluorescence and diffuse reflectance spectroscopy and spectral imaging for breast surgical margin analysis. Lasers Surg Med. 2010 Jan 1;42(1):15–23.

95. Nguyen FT, Zysk AM, Chaney EJ, Kotynek JG, Oliphant UJ, Bellafiore FJ, et al. Intraoperative evaluation of breast tumor margins with optical coherence tomography. Cancer Res. 2009;69(22):8790–6.

96. Allen WM, Chin L, Wijesinghe P, Kirk RW, Latham B, Sampson DD, et al. Wide-field optical coherence micro-elastography for intraoperative assessment of human breast cancer margins. Biomed Opt Express. 2016;7(10):4139–53.

97. Erickson-Bhatt SJ, Nolan RM, Shemonski ND, Adie SG, Putney J, Darga D, et al. Real-time imaging of the resection bed using a handheld probe to reduce incidence of microscopic positive margins in cancer surgery. Cancer Res. 2015 Sep;75(18):3706–12.

98. Koya SK, Brusatori M, Yurgelevic S, Huang C, Werner CW, Kast RE, et al. Accurate identification of breast cancer margins in microenvironments of ex-vivo basal and luminal breast cancer tissues using Raman spectroscopy. Prostaglandins Other Lipid Mediat. 2020;151:106475.

99. Allweis TM, Kaufman Z, Lelcuk S, Pappo I, Karni T, Schneebaum S, et al. A prospective, randomized, controlled, multicenter study of a real-time, intraoperative probe for positive margin detection in breast-conserving surgery. Am J Surg. 2008;196(4):483–9.

100. Schnabel F, Boolbol SK, Gittleman M, Karni T, Tafra L, Feldman S, et al. A randomized prospective study of lumpectomy margin assessment with use of marginprobe in patients with nonpalpable breast malignancies. Ann Surg Oncol. 2014;21(5):1589–95.

101. Greaves J, Roboz J. Definitions Concerning Instruments, Mass, m/z, and Ions. In: Mass Spectrometry for the Novice. 1st ed. Boca Raton: Taylor & Francis; 2014. p. 2–11.

102. Zhang J, Rector J, Lin JQ, Young JH, Sans M, Katta N, et al. Nondestructive tissue analysis for ex vivo and in vivo cancer diagnosis using a handheld

mass spectrometry system. Sci Transl Med. 2017 Sep;9(406):1–11.

103. Sans M, Zhang J, Lin JQ, Feider CL, Giese N, Breen MT, et al. Performance of the MasSpec Pen for Rapid Diagnosis of Ovarian Cancer. Clin Chem. 2019/02/15. 2019 May;65(5):674–83.

104. Zhang J, Sans M, DeHoog RJ, Garza KY, King ME, Feider CL, et al. Clinical Translation and Evaluation of a Handheld and Biocompatible Mass Spectrometry Probe for Surgical Use. Clin Chem. 2021 Jul 15;

105. Keating MF, Zhang J, Feider CL, Retailleau S, Reid R, Antaris A, et al. Integrating the MasSpec Pen to the da Vinci Surgical System for In Vivo Tissue Analysis during a Robotic Assisted Porcine Surgery. Anal Chem. 2020 Sep 1;92(17):11535–42.

106. Schäfer K-C, Szaniszló T, Günther S, Balog J, Dénes J, Keserű M, et al. In Situ, Real-Time Identification of Biological Tissues by Ultraviolet and Infrared Laser Desorption Ionization Mass Spectrometry. Anal Chem. 2011 Mar 1;83(5):1632–40.

107. Fatou B, Saudemont P, Leblanc E, Vinatier D, Mesdag V, Wisztorski M, et al. In vivo Real-Time Mass Spectrometry for Guided Surgery Application. Sci Rep. 2016 May 18;6(25919).

108. Woolman M, Gribble A, Bluemke E, Zou J, Ventura M, Bernards N, et al. Optimized Mass Spectrometry Analysis Workflow with Polarimetric Guidance for ex vivo and in situ Sampling of Biological Tissues. Sci Rep. 2017 Mar 28;7(468).

109. Ogrinc N, Saudemont P, Balog J, Robin Y-M, Gimeno J-P, Pascal Q, et al. Water-assisted laser desorption/ionization mass spectrometry for minimally invasive in vivo and real-time surface analysis using SpiderMass. Nat Protoc. 2019;14(11):3162–82.

110. Saudemont P, Quanico J, Robin Y-M, Baud A, Balog J, Fatou B, et al. Real-Time Molecular Diagnosis of Tumors Using Water-Assisted Laser Desorption/Ionization Mass Spectrometry Technology. Cancer Cell. 2018;34(5):840–51.

111. Woolman M, Kuzan-Fischer CM, Ferry I, Kiyota T, Luu B, Wu M, et al. Picosecond Infrared Laser Desorption Mass Spectrometry Identifies Medulloblastoma Subgroups on Intrasurgical Timescales. Cancer Res. 2019 May 1;79(9):2426–34.

112. Schäfer K-C, Dénes J, Albrecht K, Szaniszló T, Balog J, Skoumal R, et al. In Vivo, In Situ Tissue Analysis Using Rapid Evaporative Ionization Mass Spectrometry. Angew Chemie Int Ed. 2009;48(44):8240–2.

113. Mason TJ, Bettenhausen HM, Chaparro JM, Uchanski ME, Prenni JE. Evaluation of ambient mass spectrometry tools for assessing inherent postharvest pepper quality. Hortic Res. 2021;8(1):160.

114. Ross A, Brunius C, Chevallier O, Dervilly G, Elliott C, Guitton Y, et al. Making complex measurements of meat composition fast: Application of rapid evaporative ionisation mass spectrometry to measuring meat quality and fraud. Meat Sci. 2021;181:108333.

115. Vaysse P-M, Grabsch HI, van den Hout MFCM, Bemelmans MHA, Heeren RMA, Olde Damink SWM, et al. Real-time lipid patterns to classify viable and necrotic liver tumors. Lab Investig. 2021;101(3):381–95.

116. Janssen NNY, Kaufmann M, Santilli A, Jamzad A, Vanderbeck K, Ren KYM, et al. Navigated tissue characterization during skin cancer surgery. Int J Comput Assist Radiol Surg. 2020;15(10):1665–72.

117. Tzafetas M, Mitra A, Paraskevaidi M, Bodai Z, Kalliala I, Bowden S, et al. The intelligent knife (iKnife) and its intraoperative diagnostic advantage for the treatment of cervical disease. Proc Natl Acad Sci. 2020 Mar 31;117(13):7338–46.

118. Mason S, Manoli E, Poynter L, Alexander J, Paizs P, Adebesin A, et al. Mass spectrometry transanal minimally invasive surgery (MS-TAMIS) to promote organ preservation in rectal cancer. Surg Endosc. 2020 Aug;34(8):3618–25.

119. St John ER, Balog J, McKenzie JS, Rossi M, Covington A, Muirhead L, et al. 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. 2017;19(59).

120. Balog J, Sasi-Szabó L, Kinross J, Lewis MR, Muirhead LJ, Veselkov K, et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. Sci Transl Med. 2013;5(194):1–13.

121. Santilli AML, Jamzad A, Sedghi A, Kaufmann M, Logan K, Wallis J, et al. Domain adaptation and self-supervised learning for surgical margin detection. Int J Comput Assist Radiol Surg. 2021;16(5):861–9.

122. Greaves J, Roboz J. Buying a Mass Spectrometer. In: Mass Spectrometry for the Novice. 1st ed. Boca Raton: Taylor & Francis; 2014. p. 198–203.

123. Greaves J, Roboz J. Measures of Instrument Performance. In: Mass Spectrometry for the Novice. 1st ed. Boca Raton: Taylor & Francis; 2014. p. 111–22.

124. Casey VJ, Martin C, Curtin P, Buckley K, McNamara LM. Comparison of

Surgical Smoke Generated During Electrosurgery with Aerosolized Particulates from Ultrasonic and High-Speed Cutting. Ann Biomed Eng. 2021;49(2):560–72.

125. Manoli E, Mason S, Ford L, Adebesin A, Bodai Z, Darzi A, et al. Validation of Ultrasonic Harmonic Scalpel for Real-Time Tissue Identification Using Rapid Evaporative Ionization Mass Spectrometry. Anal Chem. 2021 Apr 13;93(14):5906–16.

126. Sukhikh G, Chagovets V, Wang X, Rodionov V, Kometova V, Tokareva A, et al. Combination of Low-Temperature Electrosurgical Unit and Extractive Electrospray Ionization Mass Spectrometry for Molecular Profiling and Classification of Tissues. Molecules. 2019;24(16):1–10.

127. Sran H, Sebastian J, Hossain MA. Electrosurgical devices: are we closer to finding the ideal appliance? A critical review of current evidence for the use of electrosurgical devices in general surgery. Expert Rev Med Devices. 2016 Feb 1;13(2):203–15.

128. Munro MG. Fundamentals of Electrosurgery Part I: Principles of Radiofrequency Energy for Surgery BT - The SAGES Manual on the Fundamental Use of Surgical Energy (FUSE). In: Feldman L, Fuchshuber P, Jones DB, editors. New York, NY: Springer New York; 2012. p. 15–59.

129. El-Sayed M, Mohamed S, Saridogan E. Safe use of electrosurgery in gynaecological laparoscopic surgery. Obstet Gynaecol. 2020 Jan 1;22(1):9–20.

130. St John ER, Balog J, McKenzie JS, Rossi M, Covington A, Muirhead L, et al. _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. 2017 May;19(59).

131. Giordano BP. Don’t be a victim of surgical smoke. AORN J. 1996 Mar 1;63(3):520–2.

132. Karjalainen M, Kontunen A, Saari S, Rönkkö T, Lekkala J, Roine A, et al. The characterization of surgical smoke from various tissues and its implications for occupational safety. PLoS One. 2018 Apr 12;13(4):1–13.

133. Ulmer BC. The Hazards of Surgical Smoke. AORN J. 2008 Oct 18;87(4):721–38.

134. Choi DH, Choi SH, Kang DH. Influence of Surgical Smoke on Indoor Air Quality in Hospital Operating Rooms. Aerosol Air Qual Res. 2017;17(3):821–30.

135. Choi SH, Kwon TG, Chung SK, Kim T-H. Surgical smoke may be a

biohazard to surgeons performing laparoscopic surgery. Surg Endosc. 2014;28(8):2374–80.

136. Cheng M-H, Chiu C-H, Chen C-T, Chou H-H, Pao L-H, Wan G-H. Sources and components of volatile organic compounds in breast surgery operating rooms. Ecotoxicol Environ Saf. 2021;209(111855).

137. Limchantra I V, Fong Y, Melstrom KA. Surgical Smoke Exposure in Operating Room Personnel: A Review. JAMA Surg. 2019 Oct 1;154(10):960–7.

138. Eiceman GA, Karpas Z, Hill HH. Background. In: Ion Mobility Spectrometry. 3rd ed. Boca Raton: CRC press; 2013. p. 1–7.

139. Shvartsburg AA. IMS: Ion Dynamics and Consequent General Features. In: Differential ion mobility spectrometry : nonlinear ion transport and fundamentals of FAIMS. 1st ed. Boca Raton: CRC press; 2009. p. 4–32.

140. Dodds JN, May JC, McLean JA. Correlating Resolving Power, Resolution, and Collision Cross Section: Unifying Cross-Platform Assessment of Separation Efficiency in Ion Mobility Spectrometry. Anal Chem. 2017 Nov 21;89(22):12176–84.

141. Buryakov IA, Krylov E V, Makas EG, Nazarov EG, Pervukhin V V, Rasulev UK. Ion division by their mobility in high-tension alternating electric field. Pisma v Zhurnal Tekhnicheskoi Fiz. 1991;17(12):60–5.

142. Eiceman GA, Karpas Z, Hill HH. Ion Sources. In: Ion Mobility Spectrometry. 3rd ed. Boca Raton: CRC press; 2013. p. 71–89.

143. House JE. Acid–Base Chemistry. In: Inorganic Chemistry. 2nd ed. Elsevier; 2013. p. 273–312.

144. Eiceman GA, Karpas Z, Hill HH. Semivolatile Samples. In: Ion Mobility Spectrometry. 3rd ed. Boca Raton: CRC press; 2013. p. 49–59.

145. Safaei Z, Eiceman GA, Puton J, Stone JA, Nasirikheirabadi M, Anttalainen O, et al. Differential Mobility Spectrometry of Ketones in Air at Extreme Levels of Moisture. Sci Rep. 2019;9(5593).

146. Eiceman GA, Karpas Z, Hill HH. Effects of Analyte Concentrations. In: Ion Mobility Spectrometry. 3rd ed. Boca Raton: CRC press; 2013. p. 260–7.

147. Kolakowski BM, Mester Z. Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS). Analyst. 2007;132(9):842–64.

148. Anttalainen O, Puton J, Kontunen A, Karjalainen M, Kumpulainen P, Oksala N, et al. Possible strategy to use differential mobility spectrometry in

real time applications. Int J Ion Mobil Spectrom. 2019;

149. Eiceman GA, Karpas Z, Hill HH. Motion of Slow Ions in Gases. In: Ion Mobility Spectrometry. 3rd ed. Boca Raton: CRC press; 2013. p. 216–9.

150. Krylov E V, Nazarov EG. Electric field dependence of the ion mobility. Int J Mass Spectrom. 2009;285(3):149–56.

151. Eiceman GA, Karpas Z, Hill HH. Differential Mobility Spectrometer and the Dependence of Ion Mobility on the Electric Field Strength. In: Ion Mobility Spectrometry. 3rd ed. Boca Raton: CRC press; 2013. p. 229–35.

152. Krylov E V, Nazarov EG, Miller RA. Differential mobility spectrometer: Model of operation. Int J Mass Spectrom. 2007;266(1):76–85.

153. Salehi B, Upadhyay S, Erdogan Orhan I, Kumar Jugran A, L D Jayaweera S, A Dias D, et al. Therapeutic Potential of α- and β-Pinene: A Miracle Gift of Nature. Biomolecules. 2019 Nov 14;9(11):738.

154. Limero T, Cheng P, Boyd J. Evaluation of Gas Chromatography-Differential Mobility Spectrometry for Measurement of Air Contaminants in Spacecraft. In: International Conference On Environmental Systems. SAE International; 2006. p. 1–11.

155. Hokkinen L, Kesti A, Lepomäki J, Anttalainen O, Kontunen A, Karjalainen M, et al. Differential mobility spectrometry classification of bacteria. Future Microbiol. 2020 Mar 1;15(4):233–40.

156. Arasaradnam RP, McFarlane M, Daulton E, Skinner J, O'Connell N, Wurie S, et al. Non-invasive exhaled volatile organic biomarker analysis to detect inflammatory bowel disease (IBD). Dig Liver Dis. 2016;48(2):148–53.

157. Anttalainen A, Mäkelä M, Kumpulainen P, Vehkaoja A, Anttalainen O, Oksala N, et al. Predicting lecithin concentration from differential mobility spectrometry measurements with linear regression models and neural networks. Talanta. 2021;225:121926.

158. Shvartsburg AA. Relating IMS Data to Molecular Structure. In: Differential ion mobility spectrometry : nonlinear ion transport and fundamentals of FAIMS. 1st ed. Boca Raton: CRC press; 2009. p. 33–54.

159. Shvartsburg AA. Separations in Heteromolecular Media. In: Differential ion mobility spectrometry : nonlinear ion transport and fundamentals of FAIMS. 1st ed. Boca Raton: CRC press; 2009. p. 174–86.

160. Shvartsburg AA, Isaac G, Leveque N, Smith RD, Metz TO. Separation and Classification of Lipids Using Differential Ion Mobility Spectrometry. J Am Soc Mass Spectrom. 2011 May 10;22(7).

161. Hancock SE, Poad BLJ, Willcox MDP, Blanksby SJ, Mitchell TW. Analytical separations for lipids in complex, nonpolar lipidomes using differential mobility spectrometry. J Lipid Res. 2019 Nov 1;60(11):1968–78.

162. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods. 2018;15(4):233–4.

163. Sammut C, Webb GI. Unsupervised Learning. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 1304.

164. Sammut C, Webb GI. Supervised Learning. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 1213–4.

165. Duda RO, Hart PE, Stork DG. Fisher Linear Discriminant. In: Pattern classification. 2nd ed. New York, NY: John Wiley and Sons Inc.; 2001. p. 44–7.

166. Duda RO, Hart PE, Stork DG. Multiple Discriminant Analysis. In: Pattern classification. 2nd ed. New York, NY: John Wiley and Sons Inc.; 2001. p. 47–51.

167. Wang S, Tang J, Liu H. Feature Selection. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 503–11.

168. Marcano-Cedeño A, Quintanilla-Domínguez J, Cortina-Januchs MG, Andina D. Feature selection using Sequential Forward Selection and classification applying Artificial Metaplasticity Neural Network. In: IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society. 2010. p. 2845–50.

169. Narisetty NN. Bayesian model selection for high-dimensional data. In: Srinivasa Rao ASR, Rao CRBT-H of S, editors. Principles and Methods for Data Science. Elsevier; 2020. p. 207–48.

170. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. Pattern Recognit. 2018;77(5):354–77.

171. Subasi A. Supervised learning algorithms. In: Practical Machine Learning for Data Analysis Using Python. 1st ed. London: Academic Press; 2020. p. 104–90.

172. Schmidhuber J. Deep Learning. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 338–48.

173. Ting KM. Sensitivity and Specificity. In: Sammut C, Webb GI, editors.

Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 1152.

174. Trevethan R. Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. Front Public Heal. 2017;5(307).

175. Ting KM. Confusion Matrix. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 260.

176. Flach PA. ROC Analysis. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 1109–16.

177. Sammut C, Webb GI. Area Under Curve. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 61.

178. Webb GI. Overfitting. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 947–8.

179. Subasi A. Performance evaluation. In: Practical Machine Learning for Data Analysis Using Python. 1st ed. London: Academic Press; 2020. p. 14–22.

180. Justice AC, Covinsky KE, Berlin JA. Assessing the Generalizability of Prognostic Information. Ann Intern Med. 1999 Mar 16;130(6):515–24.

181. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures. Epidemiology. 2010;21(1).

182. Jeng M-J, Sharma M, Sharma L, Chao T-Y, Huang S-F, Chang L-B, et al. Raman Spectroscopy Analysis for Optical Diagnosis of Oral Cancer Detection. J Clin Med. 2019;8(9).

183. Ribitsch I, Baptista PM, Lange-Consiglio A, Melotti L, Patruno M, Jenner F, et al. Large Animal Models in Regenerative Medicine and Tissue Engineering: To Do or Not to Do. Front Bioeng Biotechnol. 2020;8(972).

184. Roth JA, Tuggle CK. Livestock models in translational medicine. ILAR J. 2015;56(1):1–6.

185. Appert-Collin J-C, Thomas D. Fibrous Media. 1st ed. Falk L, editor. Aerosol Filtration. London: ISTE Press Ltd & Elsevier Ltd; 2017. 31–47 p.

186. Kontunen A. A System for Intraoperative Cancer Margin Detection by Diathermy Smoke Analysis [Internet]. Tampere University of Technology;

2017. Available from: http://urn.fi/URN:NBN:fi:tty-201708241803

187. Thomas D, Charvet A. An Introduction to Aerosols. In: Falk L, editor. Aerosol Filtration. 1st ed. London: ISTE Press Ltd & Elsevier Ltd; 2017. p. 1–30.

188. Eiceman GA, Karpas Z, Hill HH. Mobility, Electric Field, and Pressure. In: Ion Mobility Spectrometry. 3rd ed. Boca Raton: CRC press; 2013. p. 165–7.

189. Friedman JH. Regularized Discriminant Analysis. J Am Stat Assoc. 1989 Mar 1;84(405):165–75.

190. Guo Y, Hastie T, Tibshirani R. Regularized linear discriminant analysis and its application in microarrays. Biostatistics. 2007 Jan 1;8(1):86–100.

191. Roine A, Saviauk T, Kumpulainen P, Karjalainen M, Tuokko A, Aittoniemi J, et al. Rapid and accurate detection of urinary pathogens by mobile IMS-based electronic nose: a proof-of-principle study. PLoS One. 2014;9(12).

192. Covington JA, Westenbrink EW, Ouaret N, Harbord R, Bailey C, O'Connell N, et al. Application of a Novel Tool for Diagnosing Bile Acid Diarrhoea. Sensors. 2013;13(9):11899–912.

193. Alexander J, Gildea L, Balog J, Speller A, McKenzie J, Muirhead L, et al. A novel methodology for in vivo endoscopic phenotyping of colorectal cancer based on real-time analysis of the mucosal lipidome: a prospective observational study of the iKnife. Surg Endosc. 2016/08/08. 2017 Mar;31(3):1361–70.

194. Balog J, Szaniszló T, Schaefer K-C, Denes J, Lopata A, Godorhazy L, et al. Identification of biological tissues by rapid evaporative ionization mass spectrometry. Anal Chem. 2010;82(17):7343–50.

195. Dossat N, Mangé A, Solassol J, Jacot W, Lhermitte L, Maudelonde T, et al. Comparison of supervised classification methods for protein profiling in cancer diagnosis. Cancer Inform. 2007 Jul 19;3:295–305.

196. Haapala I, Karjalainen M, Kontunen A, Vehkaoja A, Nordfors K, Haapasalo H, et al. Identifying brain tumors by differential mobility spectrometry analysis of diathermy smoke. J Neurosurg JNS. 2019;1(AOP):1–7.

197. Golf O, Strittmatter N, Karancsi T, Pringle SD, Speller AVM, Mroz A, et al. Rapid evaporative ionization mass spectrometry imaging platform for direct mapping from bulk tissue and bacterial growth media. Anal Chem. 2015;87(5):2527–34.

198. Veselkov KA, Mirnezami R, Strittmatter N, Goldin RD, Kinross J, Speller AVM, et al. Chemo-informatic strategy for imaging mass spectrometry-based hyperspectral profiling of lipid signatures in colorectal cancer. Proc

Natl Acad Sci. 2014;111(3):1216–21.

199. Mirnezami R, Spagou K, Vorkas PA, Lewis MR, Kinross J, Want E, et al. Chemical mapping of the colorectal cancer microenvironment via MALDI imaging mass spectrometry (MALDI-MSI) reveals novel cancer-associated field effects. Mol Oncol. 2014 Feb 1;8(1):39–49.

200. Jin X, Han J. K-Means Clustering. In: Sammut C, Webb GI, editors. Encyclopedia of Machine Learning and Data Mining. 2nd ed. New York, NY: Springer US; 2017. p. 695–7.

201. Karjalainen M, Kontunen A, Mäkelä M, Anttalainen O, Vehkaoja A, Oksala N, et al. Recovery characteristics of different tube materials in relation to combustion products. Int J Ion Mobil Spectrom. 2020;23(2):83–90.

202. Tehrani KA. PID Control Theory. In: Panda RC, editor. Introduction to PID Controllers Theory, Tuning and Application to Frontier Areas. 1st ed. Rijeka: IntechOpen; 2012. p. 213–28.

203. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas. 1973;33(3):613–9.

204. Falotico R, Quatto P. Fleiss' kappa statistic without paradoxes. Qual Quant. 2015;49(2):463–70.

205. Dang DS, Bastarrachea LJ, Martini S, Matarneh SK. Crystallization Behavior and Quality of Frozen Meat. Vol. 10, Foods . 2021.

206. Balog J, Perenyi D, Guallar-Hoyas C, Egri A, Pringle SD, Stead S, et al. Identification of the Species of Origin for Meat Products by Rapid Evaporative Ionization Mass Spectrometry. J Agric Food Chem. 2016;64(23):4793–800.

207. Quinn PJ, Chapman D, Keith AD. The Dynamics Of Membrane Structure. Crit Rev Biochem. 1980 Jan 1;8(1):1–117.

ORIGINAL COMMUNICATIONS

PUBLICATION I

Tissue identification in a porcine model by differential ion mobility spectrometry analysis of surgical smoke

Anton Kontunen, Markus Karjalainen, Jukka Lekkala, Antti Roine, and Niku Oksala

Annals of Biomedical Engineering, 46(8), 1091–1100.

DOI: https://doi.org/10.1007/s10439-018-2035-5

1. Title page

Tissue Identification in a Porcine Model by Differential Ion Mobility Spectrometry Analysis of Surgical Smoke

Anton Kontunen[1], Markus Karjalainen[1], Jukka Lekkala[1], Antti Roine*[2], Niku Oksala*[3,4,5]

*Authors contributed equally

[1]BioMediTech Institute and Faculty of Biomedical Sciences and Engineering, Tampere University of Technology, PL 527, 33720 Tampere, Finland, anton.kontunen@tut.fi

[2]Department of Surgery, Hatanpää Hospital, Tampere, Hatanpäänkatu 24, 33900 Tampere

[3]Division of Vascular Surgery, Tampere University Hospital; [4]Faculty of Medicine and Life Sciences, University of Tampere, Finland; [5]Finnish Cardiovascular Research Center, Tampere, Finland

Abbreviated title for the running head: Tissue identification from surgical smoke by DMS

Corresponding author: Anton Kontunen

Address: Korkeakoulunkatu 10, 33720 Tampere

Phone: 040 198 1160

Email: anton.kontunen@tut.fi

2. Abstract and key terms

Electrosurgery is widely used in various surgical operations. When tissue is cut with high-frequency current, the cell contents at the incision area evaporate and together with water and possible soot particles, form surgical smoke. The smoke contains cell metabolites, and therefore, possible biomarkers for cancer or bacterial infection. Thus, the analysis of surgical smoke could be used in intraoperative medical diagnostics. We present a method that can be used to detect the characteristics of various tissue types by means of differential ion mobility spectrometry (DMS) analysis of surgical smoke. We used our method to test tissue identification with ten different porcine tissues. We classified the DMS responses with cross-validated linear discriminant analysis models. The classification accuracy in a measurement set with ten tissue types was 95%. The presented tissue identification by DMS analysis of surgical smoke is a proof-of-concept, which opens the possibility to research the method in diagnosing human tissues and diseases in the future.

Key terms: FAIMS, electrosurgery, LDA, VOC

Abbreviations: differential ion mobility spectrometry (DMS), high-field asymmetric waveform ion mobility spectrometry (FAIMS), linear discriminant analysis (LDA), rapid evaporative ionization mass spectrometry (REIMS), mass spectrometry (MS), leave-one-out cross-validation (LOOCV), automatic tissue analysis system (ATAS), voltage amplitude of the asymmetric waveform (V_{RF}), voltage of the DC compensation field (V_C), volatile organic compound (VOC)

3. Introduction

Electrosurgery, also known as *diathermy*, is one of the most common energy-based surgical methods[2]. In electrosurgery, high-frequency (200 kHz – 3.3 MHz) alternating current (AC) is conducted to the patient by either a monopolar or bipolar electrode[10]. Depending on the properties of the AC signal, the electrosurgical instrument can either *cut* (high voltage, continuous waveform) or *coagulate* (low voltage, pulsed waveform) the tissue. Especially in the cut mode, the surgical electrode vaporizes the cell contents and produces *surgical smoke*.

Surgical smoke primarily consists of water, but the organic matter of the cells is also simultaneously evaporated[20]. This means that surgical smoke carries information about the excised tissue in the form of possible biomarkers or tissue-specific metabolites. Metabolites can be used to differentiate the type and state of tissues, as proven by Schäfer et al., who introduced a method for analyzing the surgical smoke created in electrosurgery in 2009[16]. Since then, the method has been extensively studied for numerous medical and food industry applications[3,4,18,21]. For example, in the most recent publication regarding ex vivo breast cancer identification from benign breast tissue, the method achieved 93.4% sensitivity and 94.9% specificity[18]. Although the results have been excellent in terms of diagnostic properties, the method has still not spread to clinical use. One reason for this is perhaps that the *Rapid Evaporative Ionization Mass Spectrometry* (REIMS) relies on the use of an expensive mass spectrometer (MS), which together with the *Intelligent Knife* (also known as *iKnife*) sampling system used in the studies can cost several hundred thousand dollars. Other potential factors hindering the clinical use are regulatory approvals and lack of evidence in large cohorts of heterogeneous tissues. In addition, the miniaturization of high-performance MS is challenging, making its use in an operating room problematic due to space constraints.

Despite its limitations, a system like the REIMS, which could accurately detect the excised tissue type during surgery, would be in high demand in the healthcare market. For example, in breast cancer surgeries, reoperations due to residual tumor tissue caused by errors in cancer margin detection are common and add to the total healthcare costs significantly[13]. The current gold standard for intraoperative cancer margin evaluation is histopathological examination from a frozen section. The examination is expensive, and it disrupts the flow of the operation, since the operating staff and anesthetized patient are forced to wait for the results for tens of minutes.

Besides the REIMS system, other methods that aim to help in intraoperative cancer margin detection have been introduced. One example of a commercially available device is the MarginProbe (Dune Medical Devices Ltd, Caesarea, Israel), which uses *Radio-Frequency Spectroscopy* to assess the differences in breast tissue[1]. In a recent study, the use of the MarginProbe lowered the number of reoperations by 61%[9]. However, using an additional tool during surgery means that the flow of the operation is disturbed.

To create a system that does not affect the workflow of the surgeon, we propose tissue analysis by *Ion Mobility Spectrometry* (IMS). IMS is a method similar to MS, but the main difference is that while MS requires a vacuum to operate, IMS operates in atmospheric pressure. This means that in IMS, the molecules of the medium collide

with the measured target ions[7]. Additionally, in contrast to MS, which separates molecules according to their mass and charge, IMS separates molecules based on their ion mobility, which is a combination of the electrical charge and the shape of a molecule. Furthermore, IMS devices are not only considerably cheaper and more compact than MS devices, they are also easier to produce and maintain. The main downside of IMS compared to MS is its inferior resolving power, which is due to the fact that it operates in atmospheric pressure. The resolving power can be considered as the capability of the device to differentiate overlapping ion peaks in the output spectrum.

To increase the resolving power, IMS technology has been developed further with the introduction of *Differential Ion Mobility Spectrometry* (DMS), also known as *High-field Asymmetric Waveform Ion Mobility Spectrometry* (FAIMS)[7]. In DMS, the use of a radio-frequency (RF) waveform orthogonal to the sample flow is used to further separate the sample molecules. This increases the resolving power beyond the capabilities of traditional IMS, but it is still significantly lower than in MS devices[7,22]. Despite its inferior resolving power, DMS is a tempting alternative to the cumbersome MS due to its simplicity, lower cost, and size. To our knowledge, there are no publications about the use of IMS or DMS methods to analyze surgical smoke.

In this study, we introduce a DMS-based tissue analysis system and evaluate its applicability in the discrimination of porcine tissues based on smoke created in electrosurgery. Since the availability and ethical considerations limit the use of human tissues, we considered porcine tissue as the most suitable alternative sample material for this proof-of-concept study.

4. Materials and Methods

Study Material

We obtained the tissues of a Finnish landrace pig (Sus scrofa domesticus) from a slaughterhouse (Paijan Tilateurastamo, Urjala, Finland). The tissues were slaughterhouse offal including tongue, lungs, kidneys, heart, liver, skeletal muscle, skin with subcutaneous fat, and brains. After transportation in a thermally insulated container, we stored the fresh tissue samples in a freezer at -18 °C. We froze the tissues to prevent tissue degradation during the research. We conducted the measurement phases in four weeks. Before each measurement set, we individually thawed the tissue samples and placed them on the measurement platform for analysis at room temperature. In long measurement sets (>30 min), the sample pieces were moisturized with de-ionized water using a spray bottle. The tissue samples were all from the same single animal except for the brains, since they were too small to accommodate all 60 cuts. In total 14 porcine brains were required for the measurements of gray and white matter.

Measurement Protocol

We conducted the tissue identification measurements in three phases:

Phase I: We selected five histologically distinct tissue types (skeletal muscle, fat, renal cortex, liver and lung). We performed ten electrosurgical cuts to each tissue in non-randomized order, resulting in fifty measurements. We also conducted ten ‘’empty’’ cuts, in which we turned on the knife without it touching any tissue and measured the DMS response. We added the empty cuts to the first measurement set to ensure that each tissue produced a measurable output response that could be distinguished from the baseline response. The total number of cuts in the pilot phase was chosen so that it would be possible to conduct the pilot phase in one working day. A schematic of the cutting protocol can be seen in Figure 1.

Phase II: We extended the series to ten anatomically distinct tissue types (gray and white matter of the brain, liver, skeletal muscle, subcutaneous fat, lung, renal cortex, skin, tongue and cardiac muscle). We conducted sixty cuts per tissue for a total of 600 measurements. A picture of a piece of skeletal muscle after the phase II measurements can be seen in Figure 2.

Phase III: We used the remaining tissue material from phase II to further validate the results from the previous phases by analyzing a set of four tissues (renal cortex, fat, skeletal muscle and liver). We conducted this additional small set in order to account for the possible inter-day variability of measurement conditions in the second phase. We conducted ten cuts per tissue in a randomized order for a total of forty measurements. The number of cuts was limited by the amount of available sample material left after phase II.

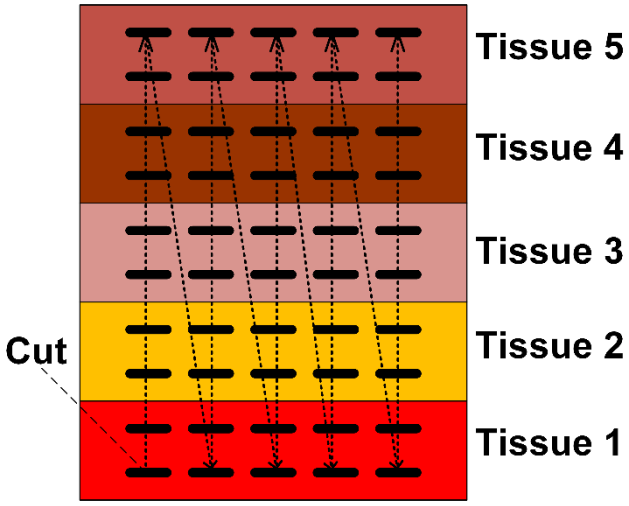


Figure 1. The order of the cuts in phase I. The empty cuts were made consecutively after the tissue cuts.

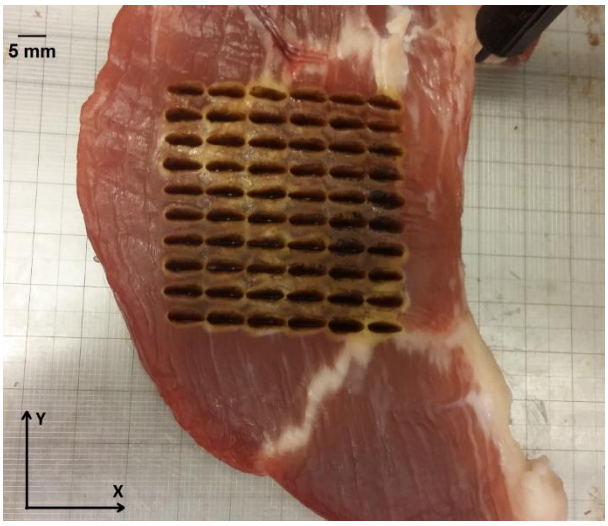


Figure 2. A piece of skeletal muscle after the measurement sequence of 60 cuts in phase II. The order of the cuts followed a similar pattern as shown in Figure 1.

The Measurement System

The measurement system we developed in this study can be divided into three distinct functional modules: automated electrosurgical sampling, sample modification (filtration and dilution), and sample detection by the ENVI-AMC® DMS device (Environics Oy, Finland). Due to the combination of the (semi-)automatic sampling system and its end-purpose, we have called our system the *automatic tissue analysis system* (ATAS).

The surgical device in the electrosurgical sampling stage of the ATAS was a medical grade system *Itkacut 350MB* (Innokas Medical Oy, Finland). The device worked with constant voltage without impedance compensation, which influenced the choice for the nominal cutting power. We selected the cutting power for the measurements with the aim of producing enough surgical smoke regardless of the tissue impedance. In the preliminary testing, it seemed that the subcutaneous fat in particular did not produce measurable concentrations of smoke with the traditionally used low cutting powers (~40 W). In order to obtain surgical smoke from all tissues, we chose 120 W

as the nominal cutting power of the surgical electrode. The electrode was a standard *monopolar knife electrode* (HF 9805-24, HEBUmedical GmbH, Germany).

The electrode was controlled by a customized *computerized numerical control* (CNC) device (REPRAP Mendel Prusa i3 kit, KitPrinter3D, Spain) that we steered with a *graphical user interface* (GUI) created with the MATLAB GUI development environment. The GUI-controlled CNC device helped in creating a standardized and stable research and development model for the measurements. Thus, the depth and duration of the electrosurgical cuts could be accurately controlled. Each cut was approximately 4 mm deep and 5 mm long, but due to the heterogeneity of the pieces of tissue, some slight variation in height was unavoidable.

To keep the height of the cuts from varying due to the position of the knife, the electrosurgical instrument was fixed to the frame of the CNC device with a polylactic acid (PLA) holder, which in turn was connected to a medical grade surgical evacuator (SURTRON® EVAC, LED SpA, Italy) by TYGON® R-3603 Laboratory Tubing (6.4×9.6 mm, Saint-Gobain, France). We used a power level of 5 for the surgical evacuator, which corresponded to an airflow of approximately 12 l/min according to measurements we performed with Gilian Gilibrator-2 NIOSH Primary Standard Air Flow Calibrator (Sensidyne, Schauenburg International GmbH, Germany). All subsequent flow rates were also measured with the Gilibrator.

Part of the airflow that went towards the surgical evacuator was guided to the particle filtration unit by a two-fold dilution system that we constructed with M/58112/09 vacuum pump ejectors (Norgen Finland Oy, Finland). We placed the ejectors to the system so that they created a negative pressure which split the sample flow before it entered the surgical evacuator. The side flow entered the filtration unit with a flow rate of 3 l/min.

The filtration unit was a newly-patented (patent pending) corona discharge filter. The corona filter was especially applicable in this research due to its ability to effectively filter out the harmful nanoscale and microscale particles in the surgical smoke (viruses, bacteria, and soot) without causing changes to the measurement dynamics of the system through pressure differences, which would be the case for example with glass fiber filters[6]. The filtration unit and the two-fold dilution system were connected to each other with Polytetrafluoroethylene (PTFE) tubing.

Even when the large particle contaminants were removed with the filtration unit, the sample concentration of the surgical smoke was often too high for the DMS sensor. Thus, we adjusted the dilution system so that the initial sample smoke was diluted with purified dry air at a ratio of approximately 1:800 before entering the ENVI-AMC® DMS device.

In the DMS device, the sample molecules first react with the so-called reactant ions created by an Americium-241 (^{241}Am) isotope and then enter a drift chamber as sample ions. In the drift chamber, the sample ions are subjected to a radio-frequency asymmetric electric field and a superimposed direct current (DC) voltage electric field[7]. Depending on the values of both the voltage amplitude of the asymmetric waveform (V_{RF}) and the voltage of the DC compensation field (V_C), sample ions with certain mobility characteristics will reach a detector plate, which creates a pA-range electric current signal upon impact. The commonly used output response of the impact spectrum, a *dispersion plot*, presents the current signals with different values for V_{RF} (y-axis) and V_C (x-axis) as a color map. In a way, the dispersion plot can be considered

as the ‘smell fingerprint’ of the measured sample. In this study, we produced the dispersion plots by increasing the value for V_C from -0.8 V to 9.0 V with 40 equal increments and by increasing the amplitude of V_{RF} from 440 V to 770 V with four equal increments. This means that the resolution for the dispersion plots was 160 (40×4) pixels. The distance between the electrodes that form the electric field in the ENVI-AMC® is 0.25 mm, so the V_{RF} values correspond to electric field strengths of 1.76 MV/m to 3.08 MV/m. Figure 3 shows example dispersion plots for liver, lung, and subcutaneous fat.

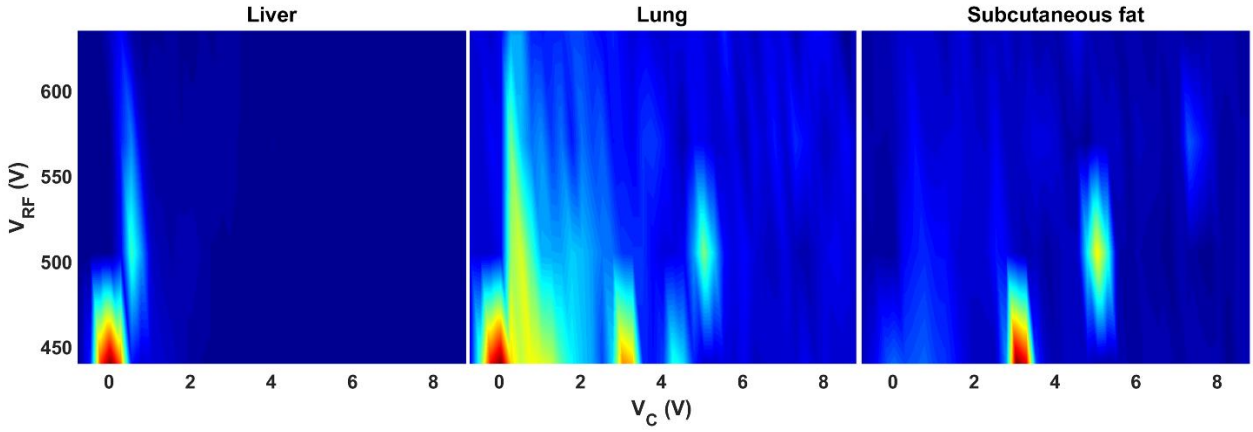


Figure 3. Example dispersion plots for liver, lung, and subcutaneous fat.

With the chosen field strengths and increments, the DMS measurement for both positive and negative ions took a total of 12 seconds per cut, after which the device cleaned itself with the dilution air for three minutes before the next measurement. This prevented any carry-over from the previous sample. A schematic representation of the measurement system and an example of an output dispersion plot from skeletal muscle tissue can be seen in Figure 4.

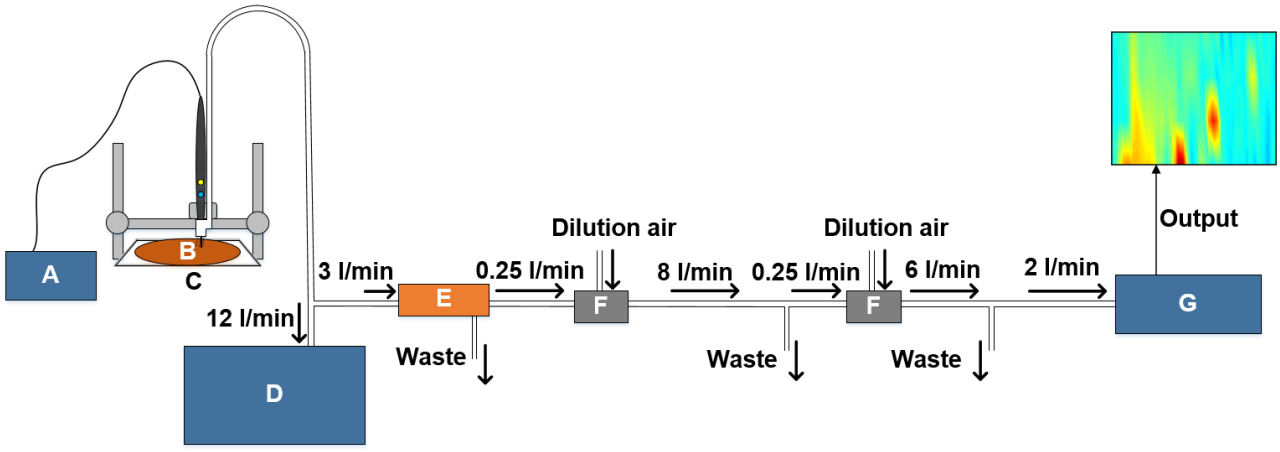


Figure 4. The measurement system. A) The Itkacut 350 MB electrosurgical unit. B) Tissue sample on the XYZ-stage of the C) CNC device. D) SURTRON® EVAC surgical evacuator. E) The filtration device. F) The dilution system. G) ENVI-AMC® differential ion mobility spectrometer.

Classification models

After each measurement set, we classified the different tissue types based on their DMS response with cross-validated *linear discriminant analysis* (LDA) models created in MATLAB (The MathWorks Inc., U.S.A). LDA is a common way to classify multi-dimensional data and it is based on reducing the dimensionality of the data by feature projection[14]. For the results in phases I and III, we cross-validated the LDA model with leave-one-out cross validation (LOOCV). For the phase II results, due to the high number of samples (600), we validated the classification model for the ten tissue types with 10-fold cross-validation. The LOOCV method is not recommended for large datasets, since the resulting training set would be almost identical with the full data set[8]. An additional validation model using 70 % of the data as training data and the remaining 30 % as test data was also done for the phase II results. This type of a *hold-out* method with 70-30 ratio for validation is a simple way to mitigate the overfitting bias in classification models, and has been commonly used in clinical studies[12].

5. Results

The results from the three measurement phases were the following:

Phase I: The five tissues were discriminated from each other and from the empty samples with a classification accuracy of 93%. One renal cortex sample was misclassified as liver and three liver samples were classified as renal cortex. All the other samples were correctly classified. The confusion matrix is presented in Table 1.

Phase II: The ten tissues were discriminated from each other with a classification accuracy of 95% by the 10-fold cross-validation model (with LOOCV, the classification accuracy was 97%). Tongue, renal cortex and skeletal muscle were most commonly misclassified. The confusion matrix is presented in Table 2. In addition to the 10-fold cross-validated results shown in Table 2, we used the phase II dataset to conduct additional validation by using 70% of the dataset to produce a classification model to discriminate the remaining 30%. The model achieved a classification accuracy of 93%. The confusion matrix for this additional model is presented in Table 3.

Phase III: The classification accuracy for the 40 measurements with four tissue types was 95%. One renal cortex sample was classified as skeletal muscle and one skeletal muscle sample was classified as liver. The confusion matrix for phase III results can be found in Table 4.

Table 1. The confusion matrix for the LDA+LOOCV classification model in phase I.[*]

Tissue	SM	SF	RC	Li	Lu	E
Skeletal muscle (SM)	10	0	0	0	0	0
Subcutaneous fat (SF)	0	10	0	0	0	0
Renal cortex (RC)	0	0	9	1	0	0
Liver (Li)	0	0	3	7	0	0
Lung (Lu)	0	0	0	0	10	0
Empty (E)	0	0	0	0	0	10

[*] LDA+LOOCV refers to leave-one-out cross-validated linear discriminant analysis. The true class of the samples is presented by the rows and the predicted class by the columns. All subsequent Tables are presented similarly.

Table 2. The confusion matrix for the 10-fold cross-validated LDA classification model in phase II.

Tissue	GM	WM	Li	SM	SF	Lu	RC	S	T	CM
Gray matter (GM)	60	0	0	0	0	0	0	0	0	0
White matter (WM)	1	59	0	0	0	0	0	0	0	0
Liver (Li)	1	0	59	0	0	0	0	0	0	0
Skeletal muscle (SM)	0	0	0	56	1	1	2	0	0	0
Subcutaneous fat (SF)	0	0	0	0	60	0	0	0	0	0
Lung (Lu)	0	0	0	0	0	55	2	0	3	0
Renal cortex (RC)	0	0	0	2	0	3	52	0	2	1
Skin (S)	0	0	0	0	0	0	1	58	1	0
Tongue (T)	0	0	0	0	0	0	6	1	52	1
Cardiac muscle (CM)	0	0	0	0	0	0	0	0	0	60

Table 3. The confusion matrix for 30% of phase II samples with the LDA classification model trained with 70 % of phase II data.

Tissue	GM	WM	Li	SM	SF	Lu	RC	S	T	CM
Gray matter (GM)	18	0	0	0	0	0	0	0	0	0
White matter (WM)	0	18	0	0	0	0	0	0	0	0
Liver (Li)	1	0	16	0	0	1	0	0	0	0
Skeletal muscle (SM)	0	0	0	16	2	0	0	0	0	0
Subcutaneous fat (SF)	0	0	0	0	18	0	0	0	0	0
Lung (Lu)	0	0	0	0	0	16	1	0	1	0
Renal cortex (RC)	0	0	0	0	0	0	16	0	2	0
Skin (S)	0	0	0	0	0	0	0	18	0	0
Tongue (T)	0	0	0	0	0	3	1	0	14	0
Cardiac muscle (CM)	0	0	0	0	0	0	0	0	0	18

Table 4. The confusion matrix for the LDA+LOOCV classification model in phase III.

Tissue	RC	SF	SM	Li
Renal cortex (RC)	9	0	1	0
Subcutaneous fat (SF)	0	10	0	0
Skeletal muscle (SM)	0	0	9	1
Liver (Li)	0	0	0	10

6. Discussion

Findings and Impact

Our results demonstrate that differential ion mobility analysis of surgical smoke can be used to differentiate healthy porcine tissues with high levels of accuracy. This finding opens up the possibility of developing an inexpensive surgical method that could be used for the near real-time assessment of tissues during electrosurgery, particularly cancer surgery. However, even though porcine tissues are a good analog for human tissues[15], in order to evaluate the clinical relevance of the technology, the findings need to be confirmed with healthy and diseased human tissues. In addition, the measurement system needs to be developed further and key ion spectrum features studied, before the method could be used in real-time analysis.

Limitations

The key limitations of our study are 1) shortcomings in the tissue specimens that limit the generalizability of the results, 2) variation in sampling and external conditions that may cause negative or positive bias to the results, 3) carry-over and the limited resolution of the dispersion plots that prevent real-time use in the current state of development.

The anatomical and mechanical variation of the tissues is also a probable cause for misclassifications. The mechanical heterogeneity of the samples with varying thickness resulted in slight variations in the cut, which changed the smoke concentration and thus, the dispersion plot. Anatomical heterogeneity is demonstrated in Figure 2 where a streak of fat is visible between muscle fibers and part of the cuts have partially hit this streak instead of pure muscle. The effect of tissue heterogeneity will be diminished as the database for the ion spectra of the tissues grows. When the properties of the ion spectra are mapped in future studies, the classification can be made by using the constant key features, which allows for more variation in the rest of the ion spectrum.

Furthermore, depending on the tissue, dispersion plots from frozen samples may differ slightly from fresh tissues. In our separate, preliminary experiments with porcine muscle tissue, there was no notable difference between the dispersion plots of fresh and frozen tissues. To generalize the results to all fresh tissues, the effect of preservation on the samples and dispersion plots needs to be explored.

Even though we standardized the electrosurgical cut, the heterogeneity of the samples led to variation in smoke concentration. This may have led to the differentiation of some tissues by the concentration of smoke rather than tissue-specific characteristics in the dispersion plots. In contrast to quantitative MS, the dispersion plot of the DMS is qualitative in nature and does not explicitly identify the detected ions or ion clusters. However, during cancer surgery, the qualitative classification of the tissue into either healthy or cancerous is sufficient for the surgeon. In the future, the concentration of smoke should be controlled to reduce bias from sampling.

The properties of the electrosurgical device may also have an effect. The device that was used in this study does not feature impedance compensation. This resulted in smoke intensity variation between the cuts, which was countered by using an

exceptionally high cutting power that might also affect the composition of the smoke, and thus the DMS response. To generalize the results to correspond with clinical use, the effect of the cutting power and impedance compensation in relation to the DMS response need to be studied further.

Additional research is also needed to find the optimal resolution for the dispersion plot. Even though the classification results were good, the DMS sweep resolution of 160 pixels is probably too low to accurately convey all the characteristic dispersion plot features of the different tissue types. The low resolution of the sweeps is a direct result of our current sampling method. In our measurements, surgical smoke was only produced during the 5 seconds of cutting, so a compromise in the resolution was needed for the DMS to match the sampling time. Low resolution is, however, only a problem in the research and development phase of the system, since the dispersion plots of the tissues are still unknown. In the future, effort should be made to increase the resolution. Higher resolution would enable researchers to more accurately study the characteristic features of the dispersion plots that are responsible for differentiating between tissues. After these features have been mapped, the DMS sweep can be concentrated on the key areas, and the measurement time will decrease to a point where it can be performed in 1–3 seconds.

Another issue that needs to be solved is carry-over. In this study, the carry-over was mitigated by a three-minute cleaning time between each cut, but a surgeon trying to distinguish tissue in the operating theatre needs a near-real-time response. This means that in order for the system to be clinically applicable, the carry-over needs to be controlled by other means.

In addition, the experimental set-up in phase II may have been subject to positive bias from the daily variations in the measurement conditions as the measurements had to be spread over several days. The absolute humidity and ambient temperature were monitored and no large shifts (>2 °C for temperature or >0.6 g/m^3 for humidity) were observed between the measurement days, but slight variations did occur. This possible variation only occurred in phase II, since the other phases were conducted in a single run of measurements during the same day.

The validation model for the phase II results (Table 3) also has potential for bias, since the measurements come from the same animal (except the brain tissues) and are spatially from a limited area. More studies are needed to conclusively prove that these results can be generalized to tissues from different pigs.

Analysis of the Results

In previous studies, porcine and other animal tissues have been identified based on surgical smoke, but the methods used have relied on complicated, expensive and bulky MS technology[5,16]. The specific molecular features that allow REIMS to differentiate porcine tissue have not been published. However, in a recent publication on the potential of using REIMS in food fraud cases[3], the method achieved 100% accuracy in determining the species that the tissue originated from. In addition, REIMS has also been used to differentiate cancerous and healthy tissue from the same organ, with above 95% sensitivity and specificity[4]. Therefore, we can assume that its performance in porcine tissue type identification will be as good as that, if not better. The classification results gained in this study are comparable and promising, when taking into account that the measurement system can still be considered an early phase prototype.

With the prototype system, the DMS responses for each tissue were different enough to enable high classification accuracy even at the low resolution of 160 pixels. This is partly explained by the different electrical properties of the tissues that lead to changes in the concentration of the smoke[17]. This phenomenon is in line with the water content of tissues[23]. However, many of the tissues (lung, liver, muscle, tongue) have similar water content, thus other factors are likely to contribute to the discrimination. REIMS studies have shown that the organic matter in the smoke from biological tissues primarily consists of various lipids (e.g. fatty acids and phospholipids) and the products of their thermal degradation[5]. These lipids are particularly common in the cell wall structures, which differ between tissue types. Even the tissues that are similar on an anatomic level such as muscle tissues (myocytes in tongue, skeletal muscle, and cardiac muscle) or neural tissues (neural cells in white matter and gray matter) exhibit different lipid profiles, which can be detected in the composition of the smoke[11].

In addition to the differences in smoke composition, and lipids, other molecules explain some of the characteristics in the dispersion plots as well as some of the misclassifications. The DMS responses for liver and renal cortex seemed to be mostly concentrated in the region of heavy volatile organic compounds (VOCs), reflected in the dispersion plots as the dominance of the ion peak at $V_C = 0$ (Figure 3). Especially in phase I, the liver and renal cortex were partly classified as each other. The misclassifications and abundant heavy VOCs can be due to metabolic waste compounds, such as bilirubin in liver and various bodily toxins in kidneys.

Aside from the liver and renal cortex in phase I, the most common misclassifications (especially in phase II) were between the tongue and renal cortex. According to chemical analysis of normal human tissues, the general composition of kidney and muscle tissue is very similar[23]. There is no comprehensive data on the molecular composition of porcine tissues. An extensive database on mouse tissue compositions did not yield findings relevant to our work[19]. The lack of relevant data on human and animal tissues with molecular profiling methods, such as MS or nuclear magnetic resonance, highlights the need for more of such studies on normal tissues.

Even though the tissue biology can partly explain the classification results, the limitations of the system and the variability of the tissue material also have to be considered. Some of the misclassifications are due to the variation within the class responses. Physical or anatomical heterogeneity and blood vessels most probably caused some tissue classes to exhibit higher variations in their ion spectrums. This, combined with the relatively low resolution, made the prediction for the classification model much more difficult.

Conclusions

We have shown that the DMS-driven differentiation of porcine tissues based on surgical smoke is possible. This study is a first step towards a novel method for surgical smoke analysis, which can foreseeably be used to discriminate between malignant and benign human tissue in the future. However, when interpreting the results, the limitations of the study need to be acknowledged. The system itself requires improvements, and additional testing using various tissue materials is needed. In addition, the method’s capacity to differentiate between diseased and healthy human tissue still has to be proven.

Almost all types of electrosurgically operated cancers can provide material for new research advancements. Further reliable and satisfactory results can eventually lead

to the commercialization of the method. The low cost and simplicity of the DMS could make the method accessible for global clinical use.

7. Acknowledgements

This study was supported by grants from the following foundations: Finnish Cultural Foundation, South Savo Regional fund, Finnish Foundation for Technology Promotion (TES), Tampereen Tuberkuloosisäätiö (Tampere Tuberculosis Foundation), Emil Aaltonen foundation and Pirkanmaan sairaanhoitopiiri (PSHP) grants 9s045, 151B03, 9T044, 9U042 and 150618.

The study material used in the research was slaughterhouse offal. Ethical approval from the Ethics Committee of the Tampere Region was not needed for conducting this research, as confirmed from the committee itself.

Authors Markus Karjalainen, Antti Roine, Niku Oksala, and Jukka Lekkala are shareholders of Olfactomics Ltd (www.olfactomics.fi), a company developing applications for eNose technology.

The authors would like to thank D.Sc. Ville Rantanen, D.Sc. Pekka Kumpulainen, and Assistant Professor Antti Vehkaoja for their comments and contributions to the final manuscript.

This is a post-peer-review, pre-copyedit version of an article published in [insert journal title]. The final authenticated version is available online at: http://dx.doi.org/10.1007/s10439-018-2035-5

8. References

1. Allweis, T. M., Z. Kaufman, S. Lelcuk, I. Pappo, T. Karni, S. Schneebaum, R. Spector, A. Schindel, D. Hershko, M. Zilberman, J. Sayfan, Y. Berlin, A. Hadary, O. Olsha, H. Paran, M. Gutman, and M. Carmon. A prospective, randomized, controlled, multicenter study of a real-time, intraoperative probe for positive margin detection in breast-conserving surgery. *Am. J. Surg.* 196:483–489, 2008.

2. Baggish, M. S., R. F. Valle, and H. Guedj. Hysteroscopy: Visual perspectives of uterine anatomy, physiology and pathology. Lippincott Williams & Wilkins, 2007, 125 pp.

3. Balog, J., D. Perenyi, C. Guallar-Hoyas, A. Egri, S. D. Pringle, S. Stead, O. P. Chevallier, C. T. Elliott, and Z. Takats. Identification of the Species of Origin for Meat Products by Rapid Evaporative Ionization Mass Spectrometry. *J. Agric. Food Chem.* 64:4793–4800, 2016.

4. Balog, J., L. Sasi-Szabó, J. Kinross, M. R. Lewis, L. J. Muirhead, K. Veselkov, R. Mirnezami, B. Dezso, L. Damjanovich, A. Darzi, J. K. Nicholson, and Z. Takáts. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. *Sci. Transl. Med.* 5:, 2013.

5. Balog, J., T. Szaniszló, K.-C. Schaefer, J. Denes, A. Lopata, L. Godorhazy, D. Szalay, L. Balogh, L. Sasi-Szabó, M. Toth, and Z. Takáts. Identification of biological tissues by rapid evaporative ionization mass spectrometry. *Anal. Chem.* 82:7343–7350, 2010.

6. Bardin-Monnier, N., and D. Thomas. 3 - Initial Pressure Drop for Fibrous Media BT - Aerosol Filtration. Elsevier, 2017, pp. 49–78.

7. Eiceman, G. A., Z. Karpas, and H. H. Hill Jr. Ion mobility spectrometry. CRC press, 2013, 1-293 pp.

8. Izenman, A. J. Recursive Partitioning and Tree-Based Methods BT - Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning. edited by A. J. Izenman. New York, NY: Springer New York, 2008, pp. 281–314.doi:10.1007/978-0-387-78189-1_9

9. Kupstas, A., W. Ibrar, R. D. Hayward, D. Ockner, C. Wesen, and J. Falk. A novel modality for intraoperative margin assessment and its impact on re-excision rates in breast conserving surgery. *Am. J. Surg.* , 2017.doi:https://doi.org/10.1016/j.amjsurg.2017.11.023

10. Munro, M. G. Fundamentals of Electrosurgery Part I: Principles of Radiofrequency Energy for Surgery BT - The SAGES Manual on the Fundamental Use of Surgical Energy (FUSE). edited by L. Feldman, P. Fuchshuber, and D. B. Jones. New York, NY: Springer New York, 2012, pp. 15–59.

11. Nikolaidis, M. G., A. Petridou, and V. Mougios. Comparison of the phospholipid

and triacylglycerol fatty acid profile of rat serum, skeletal muscle and heart. *Physiol. Res.* 55:259, 2006.

12. Pang, H., and S.-H. Jung. Sample Size Considerations of Prediction-Validation Methods in High-Dimensional Data for Survival Outcomes. *Genet. Epidemiol.* 37:276–282, 2013.

13. Pataky, R. E., and C. R. Baliski. Reoperation costs in attempted breast-conserving surgery: A decision analysis. *Curr. Oncol.* 23:314–321, 2016.

14. Quadrianto, N., and W. L. Buntine. Linear Discriminant. In: Encyclopedia of Machine Learning, edited by C. Sammut, and G. I. Webb. Boston, MA: Springer US, 2010, pp. 601–603.doi:10.1007/978-0-387-30164-8_475

15. Roth, J. A., and C. K. Tuggle. Livestock models in translational medicine. *ILAR J.* 56:1–6, 2015.

16. Schäfer, K.-C., J. Dénes, K. Albrecht, T. Szaniszló, J. Balog, R. Skoumal, M. Katona, M. Tóth, L. Balogh, and Z. Takáts. In Vivo, In Situ Tissue Analysis Using Rapid Evaporative Ionization Mass Spectrometry. *Angew. Chemie Int. Ed.* 48:8240–8242, 2009.

17. Sharp, J., K. Bouazza-Marouf, D. Noronha, and A. Gaur. Tissue type determination by impedance measurement: A bipolar and monopolar comparison. *Saudi J. Anaesth.* 11:15–20, 2017.

18. St John, E. R., J. Balog, J. S. McKenzie, M. Rossi, A. Covington, L. Muirhead, Z. Bodai, F. Rosini, A. V. M. Speller, S. Shousha, R. Ramakrishnan, A. Darzi, Z. Takats, and D. R. Leff. 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:, 2017.

19. Sugimoto, M., S. Ikeda, K. Niigata, M. Tomita, H. Sato, and T. Soga. MMMDB: mouse multiple tissue metabolome database. *Nucleic Acids Res.* 40:D809--D814, 2011.

20. Ulmer, B. C. The Hazards of Surgical Smoke. *AORN J.* 87:721–738, 2008.

21. Verplanken, K., S. Stead, R. Jandova, C. V Poucke, J. Claereboudt, J. V Bussche, S. D. Saeger, Z. Takats, J. Wauters, and L. Vanhaecke. Rapid evaporative ionization mass spectrometry for high-throughput screening in food analysis: The case of boar taint. *Talanta* 169:30–36, 2017.

22. Watson, J. T., and O. D. Sparkman. Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation. John Wiley & Sons, 2007, 25-26 pp.

23. Woodard, H. Q., and D. R. White. The composition of body tissues. *Br. J. Radiol.* 59:1209–1218, 1986.

PUBLICATION II

Identification of breast tumors from diathermy smoke by differential ion mobility spectrometry

Maiju Sutinen*, Anton Kontunen*, Markus Karjalainen, Juha Kiiski, Jill Hannus, Teemu Tolonen, Antti Roine, and Niku Oksala

*Equal contribution

European Journal of Surgical Oncology, 45(2), 141–146.

DOI: https://doi.org/10.1016/j.ejso.2018.09.005

European Journal of Surgical Oncology 45 (2019) 141–146

Contents lists available at ScienceDirect

European Journal of Surgical Oncology

journal homepage: www.ejso.com

Identification of breast tumors from diathermy smoke by differential ion mobility spectrometry☆

Maiju Sutinen [a, *, 1], Anton Kontunen [b, 1], Markus Karjalainen [b], Juha Kiiski [c], Jill Hannus [d], Teemu Tolonen [d], Antti Roine [e], Niku Oksala [a, f]

[a] Surgery, Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland
[b] BioMediTech Institute and Faculty of Biomedical Sciences and Engineering, Tampere University of Technology, Tampere, Finland
[c] Department of Musculoskeletal Disease, Division of Plastic Surgery, Tampere University Hospital, Tampere, Finland
[d] Fimlab Laboratories, Tampere, Finland
[e] Hatanpää Hospital, City of Tampere, Tampere, Finland
[f] Department of Vascular Surgery, Tampere University Hospital, Tampere, Finland

ARTICLE INFO

Article history:
Received 4 July 2018
Accepted 24 September 2018
Available online 15 October 2018

Presented at: BMT & MED Research Day 2017 held October 26, 2017 in Tampere, Finland.

Keywords:
Breast cancer
Differential ion mobility spectrometry
Surgical smoke
Surgical margin
Electrosurgery

ABSTRACT

Introduction: Breast cancer is the most frequent cancer in women worldwide. The primary treatment is breast-conserving surgery or mastectomy with an adequate clearance margin. Diathermy blade is used extensively in breast-conserving surgery. Surgical smoke produced as a side product has cancer-specific molecular features. Differential mobility spectrometry (DMS) is a rapid and affordable technology for analysis of complex gas mixtures. In our study we examined surgical smoke from malignant and benign breast tissue created with a diathermy blade using DMS.
Material and methods: Punch biopsies of 4 mm diameter from breast cancer surgical specimens were taken during gross dissection of fresh surgical specimen and placed in a well plate. The measurement system is a custom-built device called automatic tissue analysis system (ATAS) based on a DMS sensor. Each specimen was incised with a diathermy blade and the surgical smoke was analyzed.
Results: We examined 106 carcinoma samples from 21 malignant breast tumors. Benign samples (n = 198) included macroscopically normal mammary gland (n = 82), adipose tissue (n = 88) and vascular tissue (n = 28). The classification accuracy when comparing malignant samples to all benign samples was 87%. The sensitivity was 80% and the specificity was 90%. The classification accuracy of carcinomas to ductal and lobular was 94%, 47%, respectively.
Conclusions: Benign and malignant breast tissue can be identified with ATAS. These results lay foundation for intraoperative margin assessment with DMS from surgical smoke.

Introduction

Breast cancer is the most frequent cancer among women; it affects nearly 1.7 million women worldwide each year and is the fifth most common cause of cancer death in women [1]. The surgical treatment of breast cancer consists of the removal of the tumor by either breast-conserving surgery or mastectomy, accompanied with sentinel node biopsy and in select cases, removal of the axillary lymph nodes [2]. According to recent data, 60% of breast cancer is treated with breast-conserving surgery, while 80% of breast cancer patients are eligible to breast-conserving surgery, suggesting that although patient choice and logistics affect the choice of treatment, the practice may still be too radical [3,4]. Acquiring adequate margin clearance is a key challenge in breast-conserving surgery. A positive margin increases the risk of local recurrence in both invasive breast cancer and ductal carcinoma *in situ* (DCIS), warranting reoperation [5,6]. According to the Society of Surgical Oncology guidelines on margins in breast-conserving surgery, no ink on tumor consensus appears adequate for stage I and II invasive breast cancer, but for DCIS use of 2 mm margin should be the standard [7,8]. According to recent data, up to

☆ **Accepted for:** ESPRAS 2018 held 23 October – 27 October in Limassol, Cyprus.
* Corresponding author. Ratinankatu 6 A 1, FI-33100, Tampere, Finland.
E-mail address: sutinen.maiju.r@outlook.com (M. Sutinen).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ejso.2018.09.005
0748-7983/

Abbreviations

RF	Radiofrequency
OCT	Optical coherence tomography
REIMS	Rapid evaporative mass spectrometry
DMS	Differential mobility spectrometry
FAIMS	Field asymmetric ion mobility spectrometry
ATAS	Automatic tissue analysis system
LDA	Linear discriminant analysis
FSFS	Forward sequential feature selection
10-f-CV	10-fold cross validation
ROC	Receiver operating characteristic
AUC	The area under the curve

30% of patients requires reoperation due to inadequate clearance margins after breast-conserving surgery [9–11]. No-ink-on-tumor consensus has led to a decrease in reoperation rates and an increase in the popularity of breast-conserving surgery [9].

Due to costs associated with reoperations, several methods have been developed for the intraoperative evaluation of margins. An intraoperative histological examination of resected tissue by frozen section analysis can be used in suspicious areas but the entire surgical margin cannot be evaluated [12]. The diagnostic accuracy of frozen section analysis is good [13] and routine use of frozen section analysis in breast-conserving surgery has been shown to significantly reduce the reoperation rate [12,14]. Imprint cytology of specimen surfaces has been shown to have the highest performance in evaluating surgical margins of breast cancer patients [15]. The challenges of frozen section analysis include the added duration and disruption of the operation, transportation of samples and the costs of the contribution of the pathologist.

Several methods are being developed for assessing breast cancer margins intraoperatively *in vivo* or *ex vivo* besides imaging and histological assessment. Use of radiofrequency (RF) spectroscopy analysis for the detection of surgical margins (MarginProbe®) might reduce the number of reoperations due to inadequate clearance margins [16,17]. In the study by Schnabel et al. the sensitivity and specificity of the device were 75% and 46%, respectively [17]. However, RF spectroscopy disrupts the workflow of the surgeon and might compromise the orientation of the tumor leading to challenges in re-excision. Specimen radiography is used especially for nonpalpable tumors and can be a factor to consider for predicting tumor margins [18]. Intraoperative ultrasound can offer additional information on surgical margins in breast-conserving surgery [19]. Breast specimen radiography and large specimen MRI can offer additional information on the diameter of the invasive carcinoma, but for example the detection of DCIS is challenging [20]. Optical coherence tomography (OCT) has shown potential for intraoperative margin assessment in breast-conserving surgery with a sensitivity varying from 100% to 58.8% and specificity from 82% to 81% in studies [21,22].

Electrosurgery is used extensively in breast surgery [23]. It cuts tissues by heat-induced evaporation, producing surgical smoke [24]. Rapid evaporative mass spectrometry (REIMS) studies show that the molecular composition of surgical smoke has cancer-specific features with an accuracy of over 95% for discriminating malignant tissues from benign ones [25,26]. Currently, the disadvantage of mass spectrometry technology is the excessive cost and complexity, preventing adoption outside research facilities. Differential mobility spectrometry (DMS) or synonymous field asymmetric ion mobility spectrometry (FAIMS) have been extensively studied in medical applications [27]. DMS is a variant of ion mobility spectrometry: a stream of gas is first ionized, then an asymmetric radiofrequency electric field is applied to the sample gas, enabling the differentiation and filtering of ions to discriminate and analyze the molecular composition of the sample and present it as a dispersion plot. The advantage of DMS is its lower cost, complexity and maintenance requirements compared to systems based on mass spectrometry. We have demonstrated that DMS discriminates porcine tissues by surgical smoke [28]. To date, there is no data on discrimination of benign and malignant tissues.

In our study we examined surgical smoke created with a diathermy blade using DMS. We analyzed fresh malignant breast tissue, normal mammary gland and adipose tissue *ex vivo*. Our goal was to demonstrate the capability of DMS in this application, paving way for real-time margin assessment.

Material and methods

Study material

This study has been approved by the Ethical Committee of Tampere University Hospital (TAUH) (code R17007). The tissue material consisted of breast cancer surgical specimens of which surgery were performed in TAUH and Hatanpää County Hospital between June 2017 and November 2017. All patients included in the study underwent standard diagnostic workup and pre-operative tumor board. Tissue samples were collected by two breast pathologists (J.H. and T.T.) from Fimlab Laboratories, which is the principal pathology laboratory in the TAUH region.

Punch biopsies of 4 mm diameter were taken during the gross dissection of fresh surgical specimen. Tissues included in the study were: palpable breast carcinoma preliminarily assessed as clinical tumor stage (cT) $\geq$ cT2, macroscopically normal mammary gland >2 cm apart from the malignant tumor, fat and vascular structures. Several samples, usually five, were taken from each category per patient. Vascular specimens were collected only when clearly visible on grossing. The accuracy of sample site for cancerous punch biopsies were histologically verified in 20 cases (Fig. 1). The samples were covered with a gauze moisturized with saline to prevent dehydration and stored at +4 °C until analysis at Tampere University of Technology within the same or following day in most cases. For the analysis, a custom-made well plate was made to control the sampling process.

The well plate contained 40 rounded wells (5.5 mm diameter, 5.75 mm depth). To prevent the diathermy knife from short-circuiting due to direct contact with the well plate, the wells were covered with a thin protective layer of agar. Agar was a suitable material for the protective layer, since it produces unnoticeable DMS response, when subjected to electrosurgery. The surgical smoke analyses were performed in random order. Measurements of poor quality due to incomplete production of smoke, defective sample preprocessing or malfunctions in the DMS analysis were excluded from the study. The selection criteria for technically failed measurements are presented in Fig. 2. The number of measured and excluded samples for each measurement day can be seen in the supplementary material (Table S1). The full reports of the cases (structured histopathology report and a prognostic panel including estrogen receptor (ER), progesterone receptor (PgR), Ki-67, Her2 IHC and dual ISH) were gathered from patient records after analysis.

Measurement system

The measurement system used in this study is an advanced model of a custom-built device that we have previously described

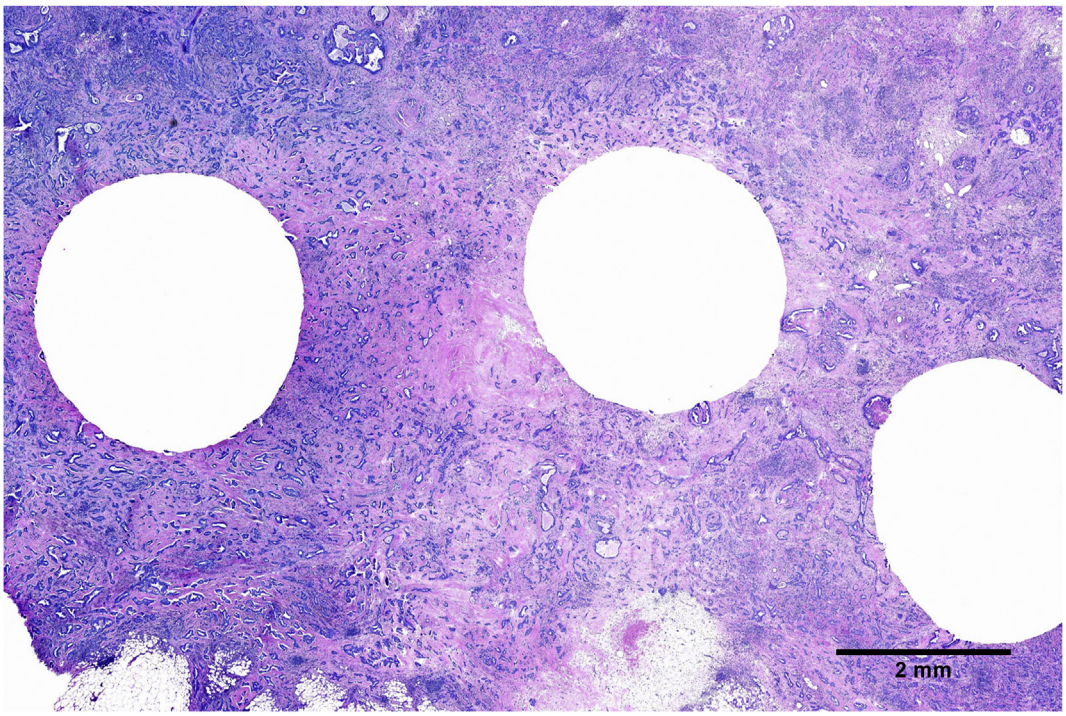


Fig. 1. Histological examination (hematoxylin and eosin staining) of cancerous tissue to verify the placement of punch biopsies taken. Scale bar presented on the lower right corner.

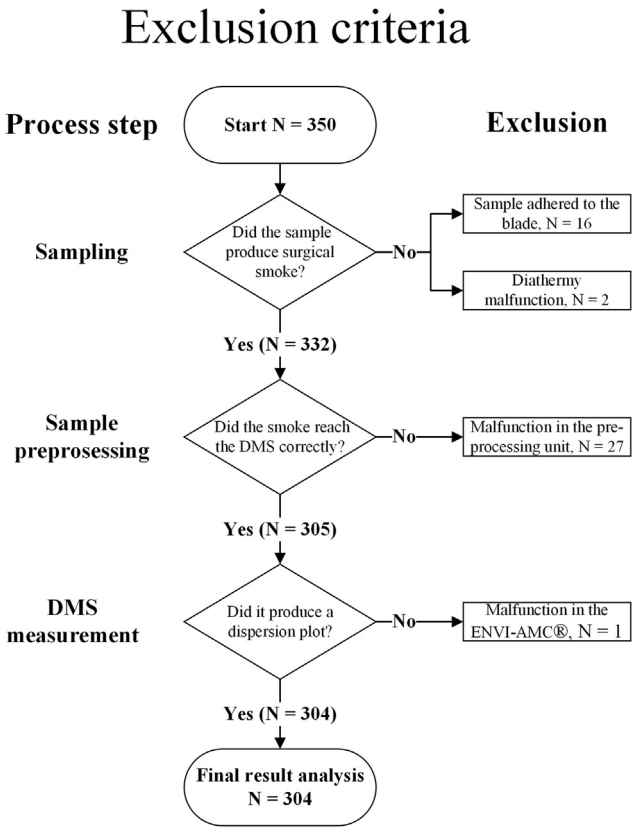


Fig. 2. Exclusion criteria of all analyzed samples.

as the *automatic tissue analysis system* (ATAS). The system comprises a computer-controlled electrosurgical sampling stage, a gas sample pre-processing unit, and the ENVI-AMC® DMS device (Environics Oy, Finland) [28].

In this study, the DMS analysis produced an output *dispersion plot* of 1620 values that represent the ion spectrum of the measured surgical smoke sample. A schematic illustration of the measurement system and an example output dispersion plot are presented in Fig. 3.

Sampling protocol

Each sample was cut individually in an automated measurement sequence. The depth (3 mm) and duration (1.5 s) of the electrosurgical cut was kept constant to stabilize the concentration of the created surgical smoke. The duration of DMS was 1 min, after which the system cleaned itself with dried purified air for another minute to prevent carry-over in the subsequent measurement.

Classification models and statistical analysis

The DMS data from the measured surgical smoke was processed with cross-validated linear discriminant analysis (LDA) classification algorithms created in MATLAB (The MathWorks Inc., U.S.A). The basic principle of LDA classification is presented in the supplementary material (Fig. S1). The LDA classification models were cross-validated with 10-fold cross validation (10-f-CV). To exclude the possible overfitting bias from the classification, an additional holdout classification with randomly selected 70% of the data used as the train set and the remaining 30% used as the test set, was also conducted.

The LDA classification of the samples was first done by using the full raw data matrix (1620 values) of the dispersion plots. The analysis was continued with a process called *forward sequential feature selection* (FSFS) to locate the key areas in the dispersion plots of the tissue types. In FSFS, subsets of the data matrices are selected until the classification accuracy does not improve significantly by the addition of new features [29]. A block diagram explaining the principle of FSFS is presented in the supplementary material (Fig. S2). The most relevant pixels in the dispersion plots were chosen for further analysis by performing 1000 cycles of FSFS each with new dataset partitions for 10-f-CV.

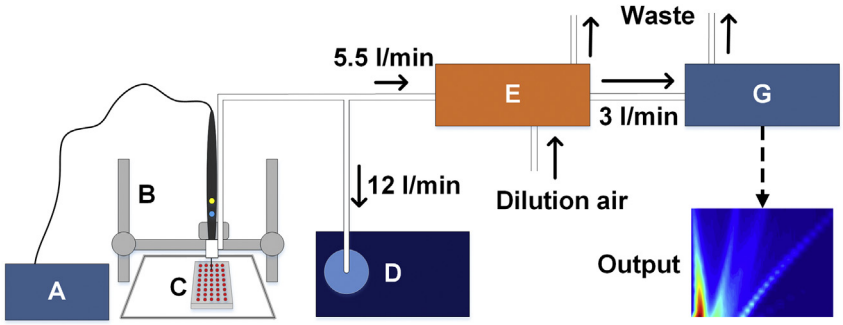


Fig. 3. The ATAS measurement system.A) The Itkacut 350 MB electrosurgical unit. B) CNC device. C) Tissue samples in the custom-made well plate. D) SURTRON® EVAC surgical evacuator. E) The sample pre-processing unit. G)ENVI-AMC® differential ion mobility spectrometer.

Using the FSFS-selected features, a receiver operating characteristic (ROC) curve was also plotted to visualize the diagnostic properties of the differentiation of carcinoma and all measured benign tissues. The ROC curve was plotted using the bootstrap method with 1000 repetitions.

Results

Characteristics of study samples

In total we examined 106 surgical smoke samples from 21 malignant breast tumors and 198 samples from benign tissues including macroscopically normal mammary gland (n = 82), adipose tissue (n = 88) and vascular tissue (n = 28). No benign breast tumors were included in the study. Clinical characteristics of carcinomas are presented in Table 1. Histopathological analysis was done for all tumors.

Table 1
Summary of clinical characteristics (n = 21).

Characteristics		Number of patients	Percentage (%)
Tumour type	Ductal	14	66.7
	Lobular	6	28.6
	Invasive micropapillary	1	4.8
Histological grade	1	5	23.8
	2	10	47.6
	3	6	28.6
Tumour size[a]	T1c (1–2 cm)	3	14.3
	T2 (2–5 cm)	14	66.7
	T3 (>5 cm)	4	19.0
Nodal status	Negative	11	52.4
	Positive	9	42.9
	Unknown	1	
HER2	Negative	18	85.7
	Positive	3	14.3
ER[b]	Negative	1	4.8
	Low	0	0
	Positive	20	95.2
PgR[c]	Negative	4	19.0
	Low	4	19.0
	Positive	13	61.9
Ki67 index	Low (≤20%)	13	61.9
	High (>20%)	8	38.1
DCIS	Present	8	38.1
	Not present	13	61.9

[a] Greatest dimension.
[b] ER denotes estrogen receptor status: positive > 20%, low 1–10% and negative < 1%.
[c] PgR denotes progesterone receptor status: positive >20%, low 1–10% and negative <1%.

Classification results

With the raw data, our 10-f-CV LDA model achieved a classification accuracy of 86.5% when comparing carcinoma samples to all benign samples. The sensitivity was 80.1% and specificity 89.9%. The results of the binary classification are presented in Table 2. With the holdout method, the classifier achieved an overall classification accuracy of 84.9%, a sensitivity of 77.1% and a specificity of 89.1%. In total, we measured 350 samples, out of which 304 were used in the final analysis. The exclusion criteria for the raw data can be seen in Fig. 2.

The classification accuracy of breast cancer to ductal (n = 69), lobular (n = 32) and invasive micropapillary (n = 5) was 94%, 47%, and 100%, respectively. The total classification accuracy by cancer type was 80.2%. The confusion matrix of the classification results is presented in the supplementary material (Table S2).

After the breast cancer classification with all 1620 pixels of the dispersion plot, the thousand FSFS cycles were performed. The process revealed that by average, 7 pixels from the dispersion plot were used in the feature selective classification of carcinoma and benign tissues. The average classification accuracy was 82.4% with a standard deviation of 1.8%. With the seven most frequently selected pixels, a ROC curve for the classification between carcinomas and benign tissue was plotted (Fig. 4). The area under the curve (AUC) value for the classification was 0.895 with 95% confidence bounds of 0.850 and 0.923.

Discussion

DMS coupled with ATAS achieves high performance in discrimination of benign and malignant breast tissue. The analysis of raw DMS data achieved a high discrimination rate of 86.5%. Similar performance was retained with only seven selected features of the dispersion plots and with the holdout validation. These results demonstrate the feasibility of intraoperative margin assessment with DMS from surgical smoke.

The inherent advantage of surgical smoke analysis is that if the surgeon uses electrocautery to excise the whole tumor, the whole

Table 2
The diagnostic table for the classification of malignant and benign breast tissues.

Positive test result	TP = 85	FP = 20	PPV ≈ **81.0 %**
Negative test result	FN = 21	TN = 178	NPV ≈ **89.4 %**
	Sensitivity ≈ **80.1 %**	Specificity ≈ **89.9 %**	

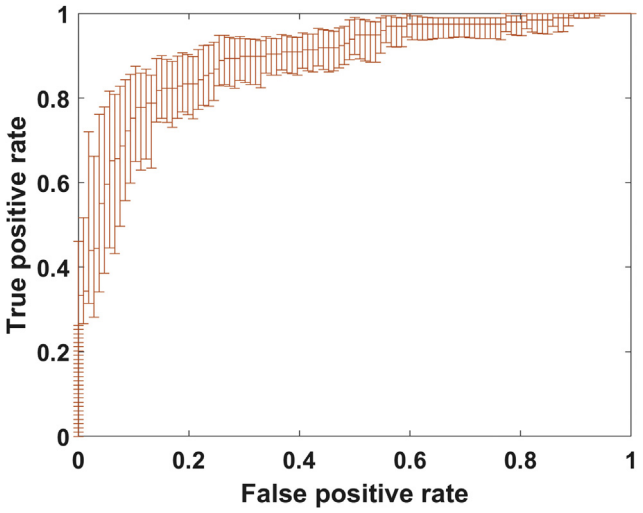


Fig. 4. ROC curve with pointwise 95% confidence bounds for the classification of malignant and benign breast tissue with FSFS-selected features.

surface of the cavity is sampled without additional stages in the surgery. Furthermore, the orientation is easy to maintain and real-time analysis would allow the surgeon to excise positive margins accurately without resorting to larger resections such as shaving the whole cavity.

The performance of ATAS is close to the reported performance of REIMS, which achieved sensitivity and specificity of 100% in a proof-of-principle study with 16 patients. The discrimination seems to result from different glycerophospholipid profiles [26]. Although the performance of REIMS is impressive, the costs and maintenance needs of mass spectrometry technology remain a major obstacle for clinical adoption. As a more economical option, DMS may achieve the optimal compromise of cost and performance.

Our results compare favorably to existing margin assessment methods. RF spectroscopy, which has been shown to reduce the re-excision rate in breast-conserving surgery in two studies [16,17], has shown comparable performance in detection of benign breast tissue and carcinoma with a sensitivity and specificity of 90% and 91%, respectively [30]. RF spectroscopy can cover an area large enough to be practical but can currently only be performed *ex vivo*, leading to challenges of accurate orientation of the resected tumor. If the margin is positive, a larger re-excision is needed to achieve negative margins because the orientation of elastic breast tissue is difficult to maintain. OCT has been applied for intraoperative breast cancer margin evaluation with a sensitivity of 100% and a specificity of 82% [21]. It yields a histologic view of the tissue, which requires human interpretation and thus significant training for the surgeon. Nolan et al. addressed this issue by developing a decision support system for studying axillary lymph nodes in breast cancer surgery with a low sensitivity of 58.8% and mediocre specificity of 81.4% [22]. The results imply that in skilled hands, the specificity of OCT is similar to ATAS. However, the results concerning the sensitivity of OCT are inconclusive. Both RF and OCT require an additional stage of analysis in the surgery, which prolongs the anesthesia and interrupts the workflow.

Our protocol was designed to minimize bias. In addition to cancerous tissue, benign samples were collected from the patients as reference. Samples were collected within hours of the operation, stored fresh and analyzed shortly after gathering. The collection of samples was done by two experienced pathologists. The analysis of surgical smoke was done in a standardized pattern in a controlled environment. Hence, the analyses done on separate occasions are comparable and the risk for day-to-day bias is low. Punch biopsies of cancerous tissue were histologically verified in 20 cases. Biopsies from one large lobular carcinoma (T3) were not verified, but due to the macroscopic extend of the carcinoma it is likely that all punch biopsies consisted of cancerous tissue. From the 20 carcinomas that were histologically verified, punch biopsies from 18 cases consisted entirely of cancerous tissue. Punch biopsies from two carcinomas were only partially cancerous tissue in the histological assessment: in total 10 punch biopsies classified as malignant tissue may contain benign tissue. When the samples from the unverified lobular carcinoma and the samples from the two carcinomas of which punch biopsies were only partially cancerous tissues were excluded from the study, the accuracy (87%), sensitivity (84%) and specificity (89%) remained unchanged suggesting insignificant bias. Therefore, the samples remained included in the study. Benign punch biopsies were not histologically verified. However, all benign punch biopsies were collected by experienced pathologists by macroscopic judgement from a distance of 2 cm from the tumor. Should some of the benign contain malignant tissue, it would cause negative bias and would underestimate the performance of DMS.

All carcinomas that were preliminarily assessed as clinical tumor stage were included in the study. Despite the preliminary estimation a portion of tumors were classified as pT1c ($n = 4$). One of the tumors was micropapillary carcinoma and consequently the number of micropapillary carcinoma samples analyzed was small compared to other carcinomas. Micropapillary and ductal carcinoma samples were analyzed on the same day: all micropapillary samples were classified correctly and only one ductal carcinoma sample was classified as invasive micropapillary. The performance of the classification could be more balanced with additional measurements to strengthen classification of carcinomas by their key features or by separating the benign classes. Weighting coefficients could also be implemented for further improving the diagnostic properties.

Even though the results of this study were comparable to previous studies, limitations of the design must be acknowledged. The key issues in the study were the challenges relating to the function of the ATAS system, variations in the study material, and the duration of the measurements. The system-related limitations led to occasional failed measurements that could not be used in the result analysis. The causes for failed measurements were 1) sample adherence to the diathermy, which prevented smoke production, 2) data communication delay between the sampling system, and 3) malfunctions of the pre-processing unit or the DMS device. Even with the exclusion of technically failed measurements, responses of especially vascular tissue specimens exhibited high variation. The heterogeneity of the tissue pieces and the resulting variation in the dispersion plots can explain the relatively low sensitivity of the binary classification. In our classification models, the benign tissue class was heterogenous with distinct dispersion plots from adipose tissue, macroscopically normal mammary gland and vascular tissue. This makes the classification criteria more complex and reduces the performance in small samples. The dispersion plots of the malignant tissues were more homogenous, leading to a more uniform malignant class. In addition, the benign class consisted of almost twice as many samples compared to the malignant class, which may bias the classification algorithm to favor the larger group in unsure cases, resulting in high specificity, but low sensitivity.

In this pilot study with the focus on the proof-of-concept, we used a high-resolution dispersion plot and a long cleaning period to maximize the data from the samples. This results in 1-min duration of the measurement with an additional 1-min for cleaning period. However, the FSFS classification models revealed that the whole

spectrum of 1620 pixels is needlessly large for accurate differentiation of carcinoma and benign breast tissues. By concentrating on the selected areas of the dispersion plot, the combined measurement and recovery time is reduced to seconds, making it feasible for real-time use. This model will be validated in future work, taking a step further towards the intraoperative cancer margin analysis.

Conclusions

The results demonstrate the ability of DMS to differentiate malignant and benign breast tissues based on surgical smoke in a laboratory setting. In the future, a rapid analysis model should be validated and the automatic tissue analysis system should be further developed for robust operation.

Declaration of interest

Niku Oksala, Antti Roine, Markus Karjalainen, and Anton Kontunen, are shareholders of Olfactomics Ltd. which is about to commercialize proprietary technology for the detection of diseases by ion mobility spectrometry.

Ethical approval

Ethical approval was obtained from the Regional Ethics Committee of Tampere University Hospital (code R17007). The study was conducted according to Declaration of Helsinki principles.

Acknowledgements

Colleagues from Tampere University of Technology and the staff of Fimlab Laboratories' pathology laboratory.

This study was supported by grants from the following foundations: Finnish Foundation for Technology Promotion (TES), Tampereen Tuberkuloosisäätiö (Tampere Tuberculosis Foundation), Emil Aaltonen foundation and Pirkanmaan sairaanhoitopiiri (PSHP) grants 9s045, 151B03, 9T044, 9U042, 150618, 9U042 and 9V044.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejso.2018.09.005.

References

[1] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136:E359–86.
[2] Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2015;26:8–30.
[3] Fajdic J, Djurovic D, Gotovac N, Hrgovic Z. Criteria and procedures for breast conserving surgery. Acta inform.med. 2013;21:16–9.
[4] DeSantis CE, Lin CC, Mariotto AB, Siegel RL, Stein KD, Kramer JL, et al. Cancer treatment and survivorship statistics, 2014. Ca - Cancer J Clin 2014;64:252–71.
[5] Houssami N, Macaskill P, Marinovich ML, Morrow M. The association of surgical margins and local recurrence in women with early-stage invasive breast cancer treated with breast-conserving therapy: a meta-analysis. Ann Surg Oncol 2014;21:717–30.
[6] Marinovich ML, Azizi L, Macaskill P, Irwig L, Morrow M, Solin LJ, et al. The association of surgical margins and local recurrence in women with ductal carcinoma in situ treated with breast-conserving therapy: a meta-analysis. Ann Surg Oncol 2016;23:3811–21.
[7] Morrow M, Van Zee KJ, Solin LJ, Houssami N, Chavez-MacGregor M, Harris JR, et al. Society of surgical oncology-american society for radiation oncology-american society of clinical Oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in ductal carcinoma in situ. [Review]. Ann Surg Oncol 2016;23:3801–10.
[8] Moran MS, Schnitt SJ, Giuliano AE, Harris JR, Khan SA, Horton J, et al. Society of Surgical Oncology-American Society for Radiation Oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in stages I and II invasive breast cancer. Int J Radiat Oncol Biol Phys 2014;88:553–64.
[9] Morrow M, Abrahamse P, Hofer TP, Ward KC, Hamilton AS, Kurian AW, et al. Trends in reoperation after initial lumpectomy for breast cancer: addressing overtreatment in surgical management. JAMA Oncol 2017;3:1352–7.
[10] McCahill LE, Single RM, Bowles EJ Aiello, Feigelson HS, James TA, Barney T, et al. Variability in reexcision following breast conservation surgery. J Am Med Assoc 2012;307:467–75.
[11] van Leeuwen MT, Falster MO, Vajdic CM, Crowe PJ, Lujic S, Klaes E, et al. Reoperation after breast-conserving surgery for cancer in Australia: statewide cohort study of linked hospital data. BMJ Open 2018;8. e020858.
[12] Jorns JM, Visscher D, Sabel M, Breslin T, Healy P, Daignaut S, et al. Intraoperative frozen section analysis of margins in breast conserving surgery significantly decreases reoperative rates: one-year experience at an ambulatory surgical center. Am J Clin Pathol 2012;138:657–69.
[13] Edward Robert St John, RBS AlKhudairi, HMRCS Ashrafian, TFECTS Athanasiou, Takats Z, DJMP Hadjiminas, et al. Diagnostic accuracy of intraoperative techniques for margin assessment in breast cancer surgery: a meta-analysis. Ann Surg 2017;265:300–10.
[14] JCFACS Boughey, TJFACS Hieken, JWFACS Jakub, ACFACS Degnim, CSFACS Grant, DRFACS Farley, et al. Impact of analysis of frozen-section margin on reoperation rates in women undergoing lumpectomy for breast cancer: evaluation of the National Surgical Quality Improvement Program data. Surgery 2014;156:190–7.
[15] Bakhshandeh M, Tutuncuoglu SO, Fischer G, Masood S. Use of imprint cytology for assessment of surgical margins in lumpectomy specimens of breast cancer patients. Diagn Cytopathol 2007;35:656–9.
[16] Allweis TM, Kaufman Z, Lelcuk S, Pappo I, Karni T, Schneebaum S, et al. A prospective, randomized, controlled, multicenter study of a real-time, intraoperative probe for positive margin detection in breast-conserving surgery. Am J Surg 2008;196:483–9.
[17] Schnabel F, Boolbol SK, Gittleman M, Karni T, Tafra L, Feldman S, et al. A randomized prospective study of lumpectomy margin assessment with use of MarginProbe in patients with nonpalpable breast malignancies. Ann Surg Oncol 2014;21:1589–95.
[18] Mazouni C, Rouzier R, Balleyguier C, Sideris L, Rochard F, Delaloge S, et al. Specimen radiography as predictor of resection margin status in non-palpable breast lesions. Clin Radiol 2006;61:789–96.
[19] Ivanovic NS, Zdravkovic DD, Skuric Z, Kostic J, Colakovic N, Stojiljkovic M, et al. Optimization of breast cancer excision by intraoperative ultrasound and marking needle - technique description and feasibility. World J Surg Oncol 2015;13:153.
[20] Abe H, Shimauchi A, Fan X, River JN, H Sattar, Mueller J, et al. Comparing postoperative human breast specimen radiograph and MRI in lesion margin and volume assessment. J Appl clin Med Phys 2012;13:3802.
[21] Nguyen FT, Zysk AM, Chaney EJ, Kotynek JG, Oliphant UJ, Bellafiore FJ, et al. Intraoperative evaluation of breast tumor margins with optical coherence tomography. Cancer Res 2009;69:8790–6.
[22] Nolan RM, Adie SG, Marjanovic M, Chaney EJ, South FA, Monroy GL, et al. Intraoperative optical coherence tomography for assessing human lymph nodes for metastatic cancer. BMC Canc 2016;16:144.
[23] Meeuwsen FC, Guedon ACP, Arkenbout EA, van der Elst M, Dankelman J, van den Dobbelsteen JJ. The art of electrosurgery: trainees and experts. Surg Innov 2017;24:373–8.
[24] Bruske-Hohlfeld I, Preissler G, Jauch K, Pitz M, Nowak D, Peters A, et al. Surgical smoke and ultrafine particles. J Occup Med Toxicol 2008;3:31.
[25] Balog J, Szaniszlo T, Schaefer K, Denes J, Lopata A, Godorhazy L, et al. Identification of biological tissues by rapid evaporative ionization mass spectrometry. Anal Chem 2010;82:7343–50.
[26] Balog J, Sasi-Szabo L, Kinross J, Lewis MR, Muirhead LJ, Veselkov K, et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. Sci Transl Med 2013;5. 194ra93.
[27] Covington JA, van der Schee MP, Edge ASL, Boyle B, Savage RS, Arasaradnam RP. The application of FAIMS gas analysis in medical diagnostics. Analyst 2015;140:6775–81.
[28] Kontunen A, Karjalainen M, Lekkala J, Roine A, Oksala N. Tissue identification in a porcine model by differential ion mobility spectrometry analysis of surgical smoke. Ann Biomed Eng 2018;48(8):1091–100.
[29] Liu H. Feature selection. In: Sammut C, Webb GI, editors. Encyclopedia of machine learning. Boston, Ma: Springer; 2011.
[30] Kaufman Z, Paran H, Haas I, Malinger P, Zehavi T, Karni T, et al. Mapping breast tissue types by miniature radio-frequency near-field spectroscopy sensor in ex-vivo freshly excised specimens. BMC Med Imaging. 2016;16:57.

PUBLICATION III

Differential mobility spectrometry imaging for pathological applications

Anton Kontunen*, Jalmari Tuominen*, Markus Karjalainen, Osmo Anttalainen, Teemu Tolonen, Pekka Kumpulainen, Maiju Lepomäki, Antti Vehkaoja, Niku Oksala, and Antti Roine

*Equal contribution

Experimental and Molecular Pathology 117, 104526.

DOI: https://doi.org/10.1016/j.yexmp.2020.104526

Experimental and Molecular Pathology 117 (2020) 104526

Contents lists available at ScienceDirect

Experimental and Molecular Pathology

journal homepage: www.elsevier.com/locate/yexmp

Differential mobility spectrometry imaging for pathological applications

Anton Kontunen[a,e,1,*], Jalmari Tuominen[b,1], Markus Karjalainen[a,e], Osmo Anttalainen[b,e], Teemu Tolonen[c], Pekka Kumpulainen[b,e], Maiju Lepomäki[b], Antti Vehkaoja[a], Niku Oksala[b,d,e], Antti Roine[b,e]

[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
[e] Olfactomics Oy, Korkeakoulunkatu 7, 33720 Tampere, Finland.

ARTICLE INFO

Keywords:
Differential ion mobility spectrometry
Diathermy
Imaging
Mass spectrometry
Tissue mapping

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.

1. 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

Abbreviations: 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

* Corresponding author.
E-mail address: anton.kontunen@tuni.fi (A. Kontunen).
[1] Authors contributed equally.

https://doi.org/10.1016/j.yexmp.2020.104526
Received 12 May 2020; Received in revised form 3 July 2020; Accepted 30 August 2020
Available online 01 September 2020
0014-4800/

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 standalone 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.

2. 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.

2.1. 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 and 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 Fig. 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 (25 × 8 = 200 pixels) for each sample point in the tissue matrix.

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

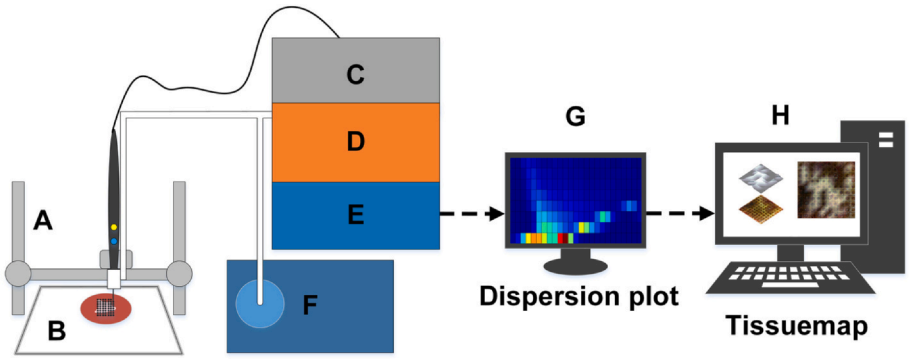


Fig. 1. The automatic tissue analysis system (ATAS) system for tissue imaging in a pathological application. A) 3D printer-based sampling unit. B) Tissue slice on the sampling platform. C) Standard diathermy unit. D) Sample pre-processing unit. E) ENVI-AMC differential mobility spectrometer (DMS). F)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

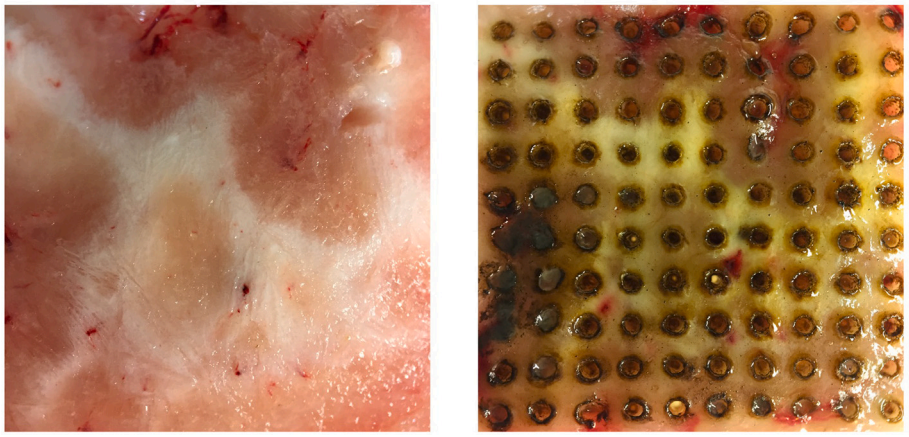


Fig. 2. Examples of images used to annotate tissues. A picture taken from a porcine brain sample before automatic tissue analysis system (ATAS) tissue mapping (left) and after the full matrix of 100 sample points with 2 mm spatial resolution for additional clarity (right).

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 s 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.

2.2. 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 (Fig. 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.

2.3. 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.

2.4. 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.

2.5. 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.

2.6. Analysis

2.6.1. 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.

2.6.2. 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.

2.6.3. 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).

3. Results

3.1. 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 Fig. 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 colour-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 Fig. 4.

3.2. 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.

3.3. 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 Fig. 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.

4. 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

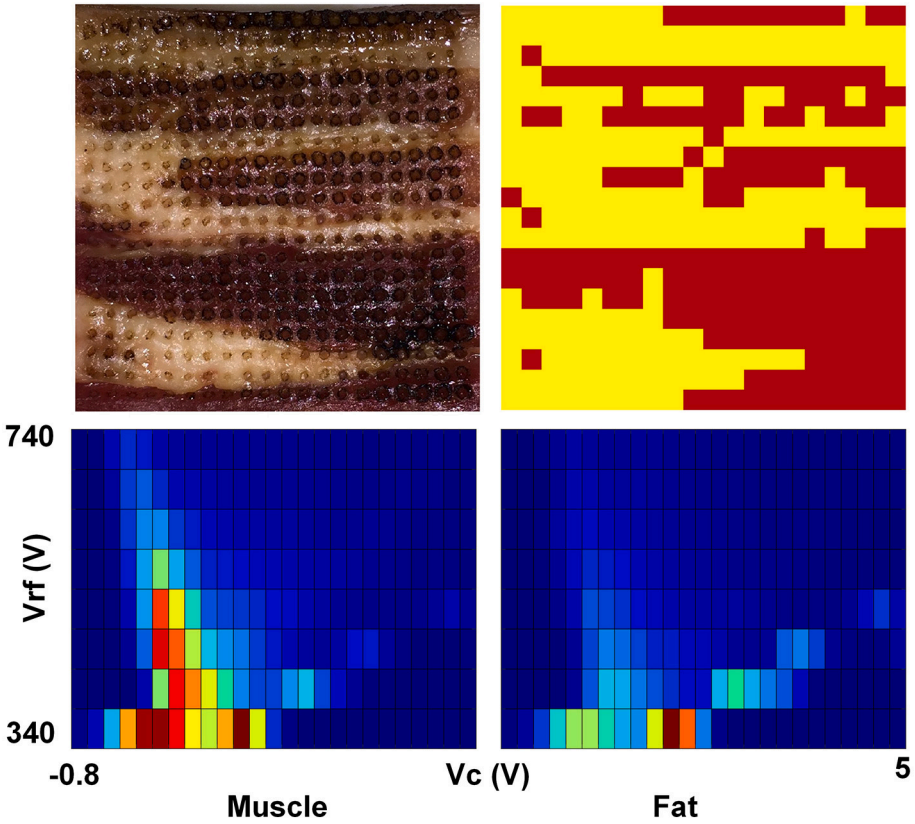


Fig. 3. K-means clustering for skeletal muscle (red) and intermuscular fat (yellow) and an example dispersion plot from both tissues. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

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.

4.1. 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

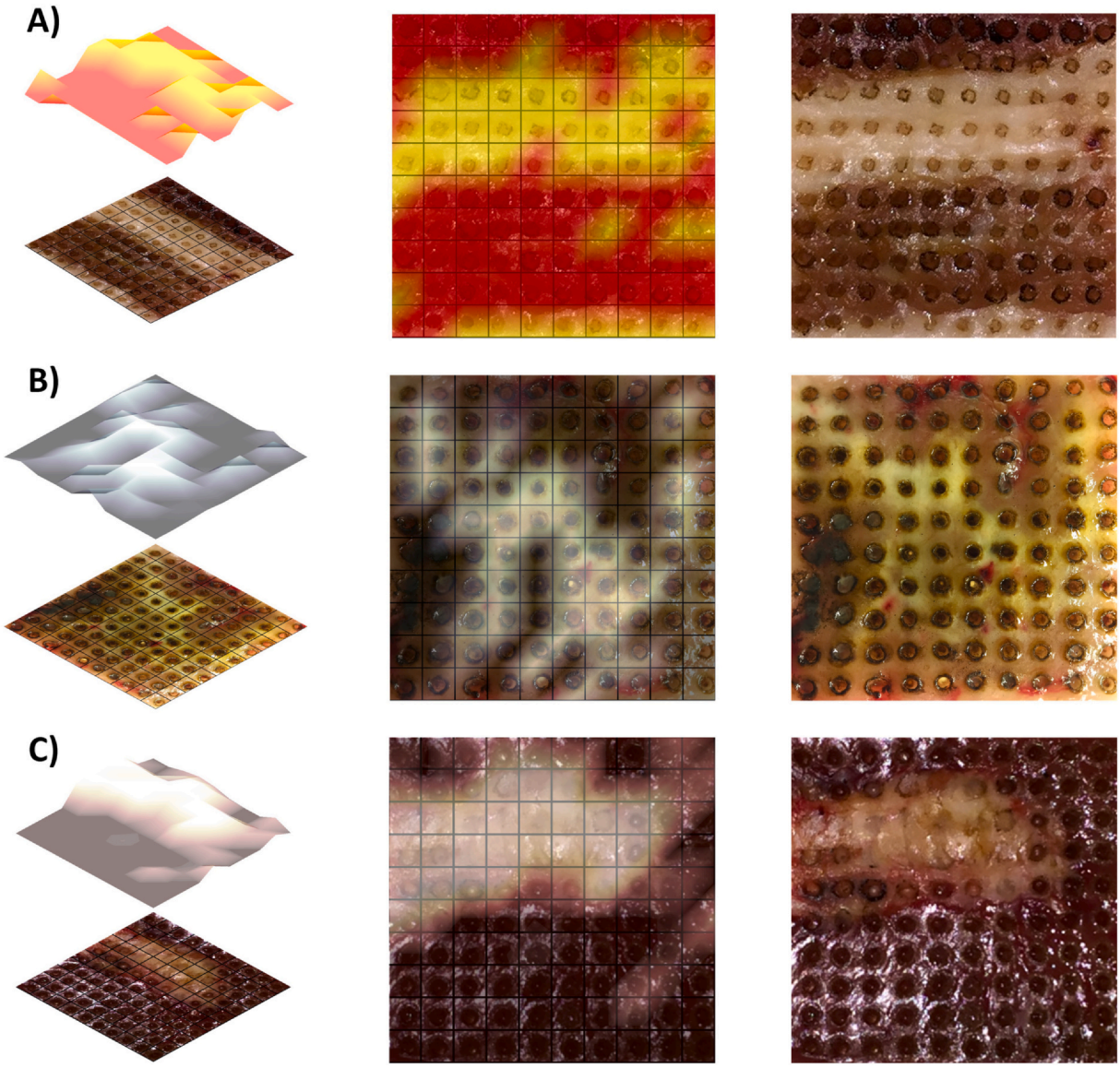


Fig. 4. Linear discriminant analysis (LDA) classification heatmap projected on top of an imaged tissue slice with 100 sample points (10 × 10 matrix with 2 mm spatial resolution for clarity). A) Porcine flank sample. The heatmap illustrates the classification certainty of each sample point. A completely 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). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

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	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
	SM	IF	RC	RP	GM	WM	Lung	Liver

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

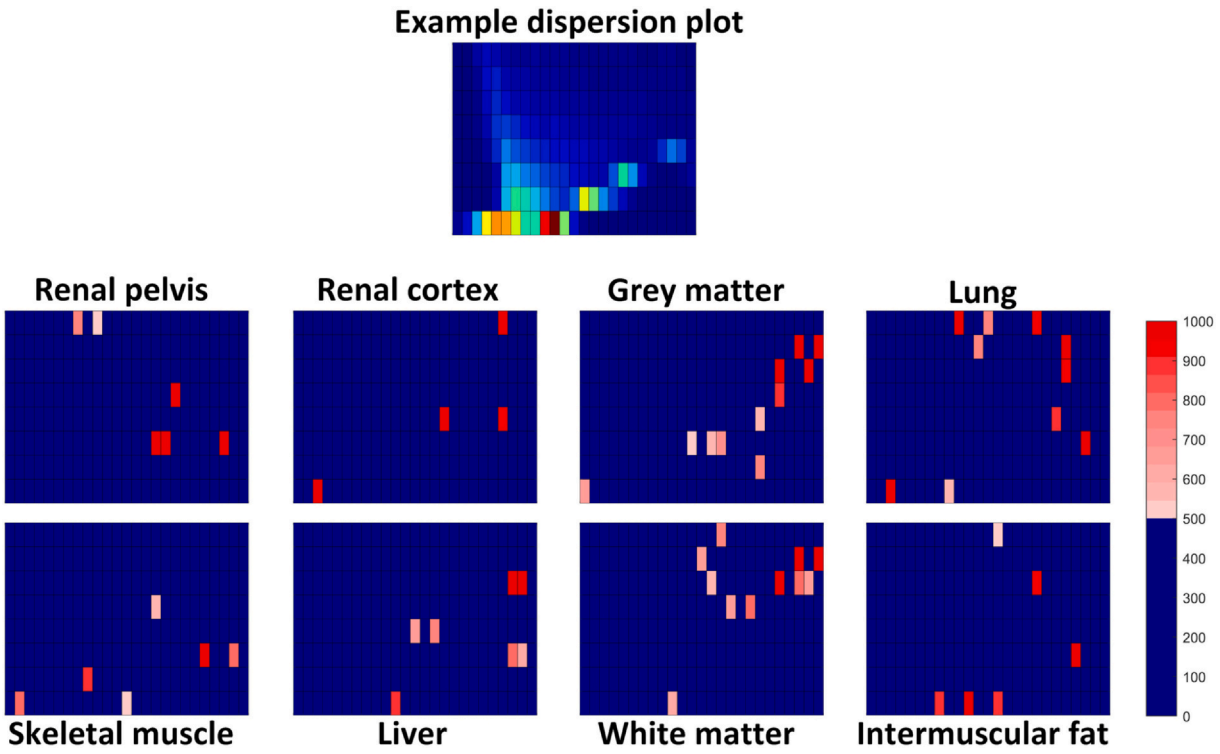


Fig. 5. Representative dispersion plot and the results of the forward sequential feature selection (FSFS) for the identification of the used tissues. The red areas of the lower dispersion plots indicate the key features that the linear discriminant analysis (LDA) classification model uses to differentiate 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. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

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.

5. 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.

Declarations

Niku Oksala (part-time employee), Antti Roine (employee), Markus Karjalainen (part-time employee), Osmo Anttalainen (employee), Pekka Kumpulainen, and Anton Kontunen (part-time employee), 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.

Funding source

Anton Kontunen declares funding from the Doctoral School of Tampere University and The Finnish Foundation for Technology Promotion. Jalmari Tuominen declares funding from The Finnish Medical Foundation and Competitive State Research Financing of the Expert Responsibility area of the Tampere University Hospital and Pirkanmaa Hospital District Grants 9X040. Maiju Lepomäki declares funding from the Doctoral School of Tampere University and The Finnish Medical Foundation. Antti Vehkaoja declares funding from Academy of Finland. Niku Oksala declares funding from Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital (9s045, 151B03, 9T044, 9U042, 150618, 9V044 and 9X040, 9AA057); from Competitive funding to strengthen university research profiles funded by Academy of Finland, decision number 292477; and from Tampereen Tuberkuloosisäätiö (Tampere Tuberculosis Foundation). The study sponsors did not have any involvement in the study design; collection, analysis and interpretation of data; the writing of the manuscript; or the decision to submit the manuscript for publication

References

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.

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/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.

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/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/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.

PUBLICATION IV

Real Time Tissue Identification From Diathermy Smoke by Differential Mobility Spectrometry

Anton Kontunen, Markus Karjalainen, Anna Anttalainen, Osmo Anttalainen, Mikko Koskenranta, Antti Vehkaoja, Niku Oksala, and Antti Roine

IEEE Sensors Journal, 21(1), 717-724.

DOI: https://doi.org/10.1109/JSEN.2020.3012965

This is an author post-print version of a published article: A. Kontunen et al., "Real Time Tissue Identification From Diathermy Smoke by Differential Mobility Spectrometry," in *IEEE Sensors Journal*, 2021, vol. 21, no. 1, pp. 717-724. Final form available online at https://doi.org/10.1109/JSEN.2020.3012965

Real Time Tissue Identification from Diathermy Smoke by Differential Mobility Spectrometry

Anton Kontunen, Markus Karjalainen, Anna Anttalainen, Osmo Anttalainen, Mikko Koskenranta, Antti Vehkaoja, Niku Oksala, and Antti Roine

Abstract

Current methods for intraoperative surgical margin assessment are inadequate in terms of diagnostic accuracy, ease-of-use, and speed of analysis. Molecular analysis of tissues could potentially overcome these issues. A system based on differential ion mobility spectrometry (DMS) analysis of surgical smoke has been proposed as one potential method, but to date, it has been able to function in a relatively slow and heavily controlled manner that is inadequate for clinical use. In this study, we present an integrated sensor system that can measure a surgical smoke sample in seconds and relay the information of the tissue type to the user in near real time in simulated surgical use. The system was validated by operating porcine adipose tissue and muscle tissue. The differentiation of these tissues based on their surgical smoke profile with a cross-validated linear discriminant analysis model produced a classification accuracy of 93.1% (N = 1059). The measurements were also classified with a convolutional neural network model, resulting in a classification accuracy of 93.2%. These results indicate that the DMS-based smoke analysis system is capable of rapid tissue identification from surgical smoke produced in freehand surgery.

1. Introduction

Surgical removal of the tumor mass is the mainstay of treatment in most solid cancers. Incomplete removal predisposes the patient to local recurrence of the tumor [1]. The goal is to remove the tumor with a margin of healthy tissue. Excessive removal of healthy tissue is associated with cosmetic or functional deficit and conservation of as much healthy tissue as possible is recommended. In the case of breast cancer, this is illustrated by the preference of *lumpectomy* (partial removal of the breast) over *mastectomy* (total removal of the breast) [2]. After the surgery, microscopic examination of the removed specimen is conducted to assess margin clearance. In 20–30% of cases,

cancerous tissue is found in or close to the tissue margin, often necessitating a reoperation, which negatively impacts patient's quality of life and yields extra costs [3]–[6]. This results in a growing need to create a tool for intraoperative tissue analysis that could reduce reoperations.

Currently, intraoperative margin monitoring is commonly conducted by imaging of the removed specimen, or by manual palpation and preoperative marking of the tumor using radiological guidance. However, certain subtypes cannot be identified by manual palpation. In these cases, microscopic analysis of frozen tissue sections or cytology smears have been shown to be effective but they are time consuming and expensive [7]–[9]. Optical coherence tomography (OCT) is a novel method that produces a microscopic view on the margin without preparation, but the interpretation of the output requires the expertise of a pathologist [10], [11].

Radiofrequency spectroscopy (RFS) detects cancerous tissue in the margin by assessing the dielectric properties of the specimen. The method has already been approved for clinical practice but has not gained wide acclaim due to poor specificity and difficulties orientating the positive finding, since the examination has to be performed *ex vivo* on removed tumor outside the patient [12].

Mass-spectrometry(MS)-based methods take tissue analysis from microscopic visual assessment to the analysis of the molecular composition of the tissues. Rapid evaporative ionization mass spectrometry (REIMS) is an experimental but promising technology that analyzes the surgical smoke that evaporates from the tissues that are cut during electrosurgical removal of the tumor [13], [14]. REIMS studies have demonstrated differences in the relative abundance of triglycerides and certain phospholipids, such as phosphatidylcholines and phosphatidylethanolamines between malignant and benign tissues. The proposed underlaying mechanism between these differences is the Warburg effect that has been well documented in various cancers [15]. REIMS is capable of differentiating benign and cancerous tissues at high accuracy within seconds during the surgery, thereby not interrupting the procedure or causing challenges in tumor orientation. Other MS-based methods include, SpiderMass, Picosecond InfraRed Laser Mass Spectrometry (PIRL-MS), and MasSpec pen [16]–[19]. SpiderMass and PIRL-MS use laser excitation to produce the gas phase sample, whereas in MasSpec pen, a water droplet is used to extract molecules from the target tissue sample. Research with these methods has been promising. Moreover, they are non-destructive by nature. However, the instruments disrupt the normal workflow of the surgeon by requiring an additional tool to be used during the surgery. All these technologies are hindered by their large physical size and the high cost of MS instrumentation.

In contrary to MS, our method is based on the usage of more affordable, almost maintenance free, differential mobility spectrometry (DMS) technology, in which gaseous compounds can be differentiated by their ion mobility characteristics in an asymmetrical electric field. Our previous research has demonstrated the potential of the DMS-based automatic tissue analysis system (ATAS) in the classification of healthy tissues, breast cancer and brain cancer [20]–[22]. In ATAS, the tissue sample is cut by controlling the electrosurgical knife with a modified 3D printer.

When the tissues are sampled with the ATAS, the electrosurgical cuts are standardized (depth, duration), thus stabilizing the smoke sample concentration [20]. However, as we intend to move towards the surgical application, the system needs to be able to function robustly with varying sampling conditions with close resemblance to a clinical situation and application. Therefore, ATAS needs to be replaced with handheld instrumentation. This in turn leads to technical requirements for the system that demand innovative engineering solutions and extensive modifications to the system compared to previously published methodology.

In this study, we introduce a tissue analysis system compliant with operation room workflow and standard

electrosurgical systems. We also describe its key engineering novelties, which include a heating solution to minimize contamination, an adaptive electric filter, a method to detect the induced diathermy current in order to time the measurement, and a surgical user interface. Furthermore, we evaluate the applicability of the system in the near real time discrimination of two porcine tissues by using machine learning methods and employ convolutional neural networks in the analysis of DMS data. The DMS-based sensor system and its technical advancements introduced in this study mark a potential breakthrough for a clinical application.

2. Materials and methods

a) Study material

The specimens used in this study were fresh commercial meat products that had distinctive areas of skeletal muscle tissue and adipose tissue as seen in Fig. 1. The use of human tissues in a proof-of-concept study is not justified due to the scarceness of the material and ethical aspects. Therefore, we chose to use porcine tissue samples (porcine flank) to demonstrate the function of the system. Another reason is that the tissues found in flank are operated in almost every electrosurgical procedure, due to them being in the way or around the target tissue, for example in cancer surgeries, which makes their identification clinically relevant.

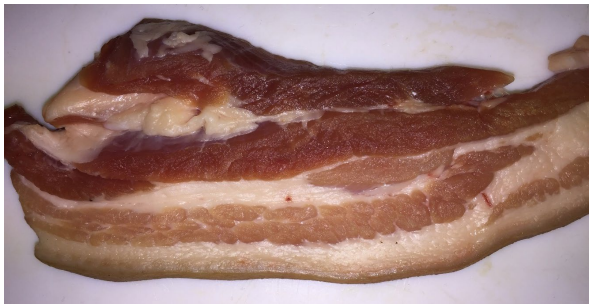


Figure 1. Porcine flank sample with macroscopically visible areas of skeletal muscle tissue and adipose tissue.

b) Measurement system and operation principle

The measurement system comprises an electrosurgical knife, surgical smoke evacuator, sample-preprocessing unit, and a DMS sensor. Fig. 2 shows a schematic representation of the measurement system with a pneumatic diagram of the sample-preprocessing unit. Fig 3. shows the measurement system in a laboratory setting.

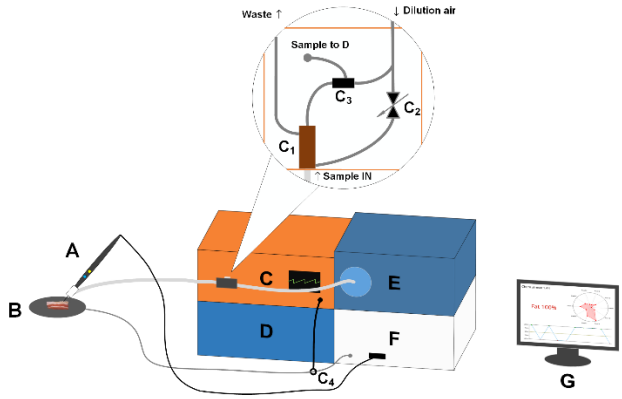


Figure 2. A stylized depiction of the measurement system with a pneumatic schematic of the sample pre-processing unit. A: Diathermy instrument with the attached 3-meter-long suction tube. B: Tissue specimen on top of the dispersive electrode. C: Sample pre-processing unit with the corona discharge filter (C_1), PID-controlled proportional valve (C_2), ejector dilution (C_3), and an inductive coil to detect the diathermy current (C_4). D: ENVI-AMC DMS sensor. E: Smoke evacuator. F: Diathermy unit. G: Surgical user interface.

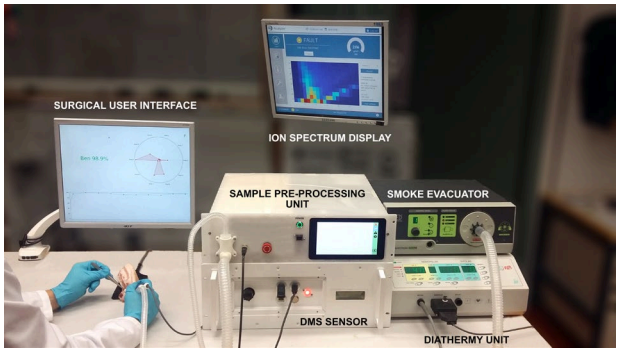


Figure 3. The measurement system in a laboratory setting.

In this study, the surgical cuts were generated by a monopolar knife electrode (A. in Fig. 2, HF 9805-24, HEBUmedical GmbH, Germany) attached to a diathermy unit (F. In Fig. 2, Itkacut 350MB, Innokas Medical Oy, Finland). The diathermy unit was set to operate at a nominal power of 40 W in 100% cut mode. The surgical smoke created during the cuts was collected by a three-meter-long suction tube (Handle aspiration kit, LED SpA, Italy) attached to the side of the diathermy instrument. From the side of the instrument, the smoke sample travelled to a T-junction that split the sample flow into two streams that led to a surgical smoke evacuator (E. in Fig. 2, SURTRON® EVAC, LED SpA, Italy) and the custom-built sample pre-processing unit (C. in Fig. 2).

The sample pre-processing unit of the current system has important additions and improvements compared to the previously described ATAS system [20], [21]. The main function of the pre-processing unit is to filter and dilute the smoke sample. However, previous studies have shown that the carry-over signal from preceding measurements is a problem when conducting rapid consecutive measurements. Therefore, the sample line in the pre-processing unit of the improved system was minimized to 75 cm to decrease the available scent adsorption area. In addition, the sample line was heated by a resistive cable heater and enclosed in a wooden box made from 4 mm thick birch plywood that keeps the sample pneumatics at elevated temperature (70°C), allowing for shorter recovery time due to reduced condensation

and scent adherence, and increased desorption of the scent molecules from the tubing. In theory, a higher temperature would work even better in mitigating the carry-over phenomenon, but 70°C was chosen as the limit to ensure that the electronic components within the pre-processing unit do not malfunction due to excessive heating.

Besides the integrated heating and minimized sample travel length, the main technical novelty of the pre-processing unit is an improved dilution system that controls the smoke sample intake based on the concentration. The adaptive dilution system is built around a filtration solution (C_1 in Fig. 2) that utilizes a corona discharge to remove particulate matter from the sample stream. A schematic representation of the filter is presented in Fig. 4. The discharge is created by applying a 5 kV voltage to the center electrode, while keeping the copper casing at ground potential. This ionizes the smoke sample and forces the particulate matter to the casing, from where it continues to a waste stream.

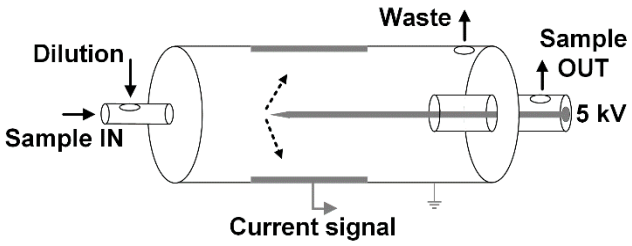


Figure 4. Corona discharge particulate filter.

In addition to removing large particles from the sample stream, the new filter measures the amount of the particulate matter. The inner wall of the filter contains electrodes, which produce a current signal, when the ionized particulate matter reaches the surface. The smoke concentration depends on how long or deep the electrosurgical cut is. The sample is mixed with dilution air. Diluted concentration is regulated with a proportional-integral-derivative (PID) feedback-controlled valve (C_2 in Fig. 2). PID is a widely used control method, where a deviation of a desired setpoint is corrected by using three corrective terms in unison to achieve minimum overshoot and delay in reaching the desired output values. In terms of the real time tissue analysis system, this means that when particulate matter causes a current signal deviation in the filter, the PID-controlled valve tries to actively compensate the change by increasing the amount of dilution air without overdiluting the sample. We have previously studied that even with similar cutting parameters, the amount of particulate matter is highly dependent on the tissue type [23]. This means that an adaptive dilution system is a useful solution to ensure that the DMS measurement does not saturate regardless of the tissue type or cutting conditions.

In addition to the adaptive dilution system, the sample pre-processing unit has a Raspberry Pi 7” Touch Screen Display (Raspberry Pi Foundation, UK) that relays information to the user about the electrical current signal of both the corona filter and the diathermy knife. The signal from the knife is monitored by an induction coil (C_4 in Fig. 2) wrapped around the cable of the dispersive electrode (B. in Fig. 2). The monitoring of the induced cutting current was utilized in the synchronization of the DMS measurement with the smoke production. In other words, the DMS measurement was triggered, when the current of the electrosurgical knife was above a fixed threshold that was exceeded only upon contact with tissue material.

The DMS sensor used in the study is an Envi-AMC® spectrometer (D. in Fig. 2) manufactured by Environics Oy, Finland. The sensor is capable of parts-per-billion-level detection of gaseous substances and operates with separation

fields between 10 to 80 Townsend (Td) with <0.01 Td resolution. The waveform is created with a field-effect transistor based half-bridge with 250 kHz operating frequency resulting in a square waveform. The waveform is run with 5 % duty cycle. The compensation field can be varied between ±1.5 Td with resolution <0.001Td. The DMS filter is arranged from three planar plates forming two identical channels, both with channel height of 0.25 mm, width of 6 mm and length of 16 mm [24]. 2.8 liters per minute sample flow rate is created with a venture nozzle, driven by clean pressurized air. Sample is drawn untreated into the DMS filter via Am-241, 5.92 MBq ionization chamber. The residence time in the ionization area is about 16 ms and in the DMS filter about 1 ms. Filtered ions are collected with three sequentially arranged electrodes. The collection field is adjusted to collect most of the ions with the middlemost and the largest of the collection electrodes. The measurement is controlled by a LUA-language script, which allows detailed control for each of the measurement parameters such as flow, frequency, duty cycle, separation field settings, compensation field settings, signal noise filtering and selection of ion polarity. The operation principle of the DMS sensor is illustrated in Fig. 5.

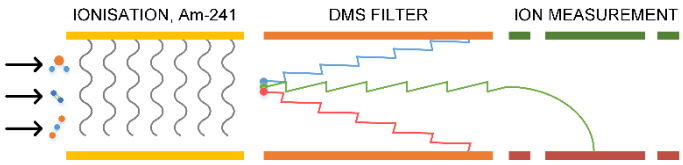


Figure 5. The operation principle of the ENVI-AMC® spectrometer. The black arrows represent the incoming gas sample molecules, and the colored lines represent sample ions that diverge in the DMS filter based on their differential mobility characteristics.

The result of a DMS measurement can be represented as a colored 2D image (i.e. dispersion plot), where the color represents the number of ions that reach the detector with different DMS filter separation and compensation voltage values. All DMS data in this paper were measured with the following settings: Positive ion mode was used, the compensation voltage sweep (x-axis in the dispersion plot) was set from -0.8 V to 5 V with 25 steps, and the separation voltage sweep (y-axis in the dispersion plot) was set from 340 V to 740 V with 8 steps. With these settings, the measurement duration was approximately 5 seconds.

After the DMS measurement of a smoke sample is finished, the data is stored to a Microsoft® Azure (Microsoft, U.S.A) cloud storage for possible offline analysis. The measurement data is also transferred to a Raspberry Pi computer (Model 3B). The computer has a linear discriminant analysis (LDA) classification model that is used to classify the newly measured data. LDA has been used as the primary classification method in our previous studies with good results and it has also been used, for example, in the REIMS studies [20], [21]. LDA classification is computationally simple and the result is produced in milliseconds. The classification script of the real time system was written in Python.

The Raspberry Pi computer that is used for online classification also displays the result in an attached monitor that functions as the surgical user interface (UI). The surgical UI (G. in Fig. 2) can be set to display the result of the LDA classification as the predicted tissue type alongside the posterior probability of the prediction. In other words, the UI can tell the surgeon how certain the model is of its prediction. This enables a more nuanced assessment of the operated tissue than a completely binary statement. In addition, the results of the previous classifications can be displayed with a graph, making following the analysis much easier.

The whole process from the tissue evaporation by freehand surgery to displaying the classification result in the surgical UI takes approximately 9 seconds. After the result is displayed, the process can start again, when the trigger level for the diathermy current exceeds the threshold level (i.e. the knife is in contact with the tissue). Fig. 6 shows the operation principle of the system in the form of a flowchart.

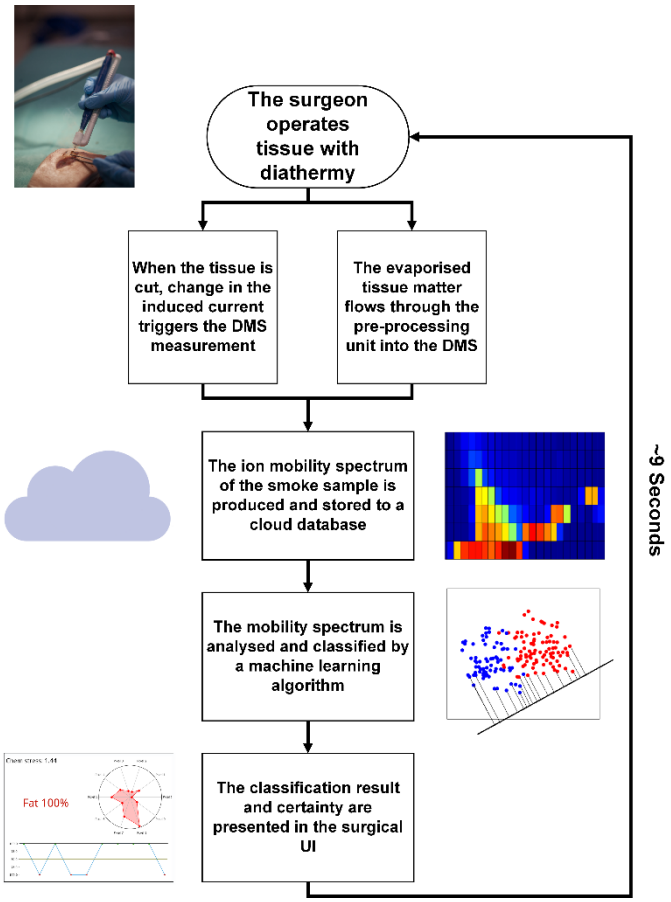


Figure 6. Flowchart of the operation principle of the real time tissue analysis system.

c) Measurement protocol and tissue sampling

The measurement protocol was designed to resemble the use of diathermy in surgical practice as closely as possible. An experienced consultant surgeon (N. Oksala) participated in the design and implementation of the protocol. The protocol consisted of three steps: 1. A fresh tissue specimen was placed on the dispersive electrode of the diathermy unit; 2. The tissue specimen was incised for 1–3 seconds with the diathermy instrument. Each incision triggered a DMS measurement of the surgical smoke. 3. A recovery period of approximately 10 seconds was taken to reduce carry-over.

The process steps 2. and 3. were repeated until no uncut surface area was available. This equated to approximately 150-200 incisions per specimen. At the end of the measurement session, the specimen was disposed, and the dispersive electrode was cleaned with isopropanol.

In total, 1159 freehand smoke samples (580 to skeletal muscle and 579 to adipose tissue) from seven specimens were measured. Out of the 1159 samples, 1059 were used for classification model training and internal validation and 100 were used for external validation. To reduce operator-dependent and temporal bias, the sample cuts were made by three operators over a period of seven measurement days.

d) Data analysis and classification

The features of the DMS data can be presented as color-coded pixels in the two-dimensional dispersion plot, where the color represents the 16-bit signal intensity value. In this respect, the DMS output dispersion plots can be considered as images that depict the chemical fingerprint of a substance, which means that image recognition methods are a viable option for the analysis and classification of DMS data. Thus, in addition to the simple LDA classification, we tested convolutional neural networks (CNN) for the classification of the tissue smoke data.

Data produced with the system were analyzed offline with MATLAB (Version R2017b, The MathWorks Inc., U.S.A). Before classification, the values of the DMS data were offset-corrected to have only positive values and normalized between 0 to 1. The offline models were implemented to have more numerical data of the performance of the classification models than what was used in the real time classification. The CNNs were created using the mxnet package for R (R Core Team, 2018). The CNNs consisted of two convolutional and two fully connected layers. The architectural and regularization parameters were chosen based on a grid search and heuristic testing. The same data was used in the hyper-parameter selection and is reported as 10-fold cross-validated (10-f-CV) results, which may have caused some positive bias in the results. To our knowledge, there are no previous publications, where CNNs have been used for DMS data.

For the performance evaluation with the training data, the models were 10-fold cross validated. In addition to the internal cross-validation, the LDA classification model was validated with an independent data set that was measured four months after the training data. The external validation with the independent test set was deemed necessary in order to assess the effect of long-term variation in the produced DMS outputs on classification performance. The classification results of the test set were analyzed both with the Raspberry embedded real time model, and offline after the sampling. The test set consisted of 100 cuts (50 skeletal muscle and 50 adipose tissue).

Further analysis of the DMS data and its features for classification were also explored by utilizing a process called forward sequential feature selection (FSFS). FSFS is an iterative process that aims to find the most significant features for the differentiation of the classes by selecting the features one-by-one until the classification performance does not significantly improve [25]. The feature selection for the LDA classification model was repeated 1000 times with different 10-fold cross-validation partitions to ensure that the selected features would not be dependent on the training set.

3. Results

The data analysis from the metadata of the DMS measurements revealed that the median time between cuts with the freehand sampling was 18 seconds (median absolute deviation 4.4 seconds). The median time is twice as long as the

minimum time for consecutive measurements, since the point of the training data was to get good representative samples of both tissues, rather than trying to minimize the analysis time.

In internal validation, the mean classification accuracy of the 10-f-CV LDA model was 93.1%, with sensitivity of 91.5% and specificity of 94.7%. The CNN model produced a mean cross-validated classification accuracy of 93.2%, with sensitivity of 91.9% and specificity of 94.5%. The results and the internal validation are presented in Table I as confusion matrices, where the true class is represented in the rows and predicted class in the columns.

In terms of external validation with the independent set, the accuracy of the real time classification was 87.0%, with sensitivity being 80.0% and specificity being 94.0%. This means that out of the 100 test samples 13 were misclassified by the real time Python model during freehand sampling. A confusion matrix of the real time classification along with the associated offline CNN classification is presented in Table II. The classification of the independent data set with the CNN model produced a 50.0% classification accuracy with all predictions being muscle tissue, i.e. sensitivity 100.0% and specificity 0%.

Due to the guess level predictions produced by the CNN model in the external validation, the training set of the model was increased with 1240 samples (721 samples of skeletal muscle and 519 samples of adipose tissue) that were not measured by freehand surgery, but instead with the ATAS system in a different study. This increased the robustness of the model and resulted in a classification accuracy of 88.0%, with sensitivity of 100.0%, and specificity of 76.0%, with the independent test set. When the additional data was also used for the training of the LDA model, the classification accuracy in offline analysis increased to 96.0%, with sensitivity of 100.0% and specificity of 92.0%. Table III shows the confusion matrices of the external validation with the CNN and LDA models trained with additional data.

The 1000 FSFS cycles with the original 1059 sample data set yielded a median number of 10 selected features from the full spectrum. By using these features for classification of the tissue types, the cross-validated mean classification accuracy was 93.2%, with sensitivity of 95.9% and specificity of 90.4%. The ten selected features that explain the differences in the spectra of muscle tissue and adipose tissue the most, are presented in Fig. 7, along with example outputs produced by both tissues.

Table 1.

Classification results with 10-fold cross validation. The true class is in rows and the predicted class in columns

	LDA			CNN	
	Fat	**Muscle**		**Fat**	**Muscle**
Fat	501	28	**Fat**	500	29
Muscle	45	485	**Muscle**	43	487

Table 2.

Classification results with the independent external test set. The true class is in rows and the predicted class in columns

	LDA			CNN	
	Fat	**Muscle**		**Fat**	**Muscle**
Fat	47	3	**Fat**	0	50
Muscle	10	40	**Muscle**	0	50

Table 3.

Classification results with the independent external test set with models that were trained with additional data. True class is in rows and predicted class in columns

	LDA			CNN	
	Fat	Muscle		Fat	Muscle
Fat	46	4	Fat	38	12
Muscle	0	50	Muscle	0	50

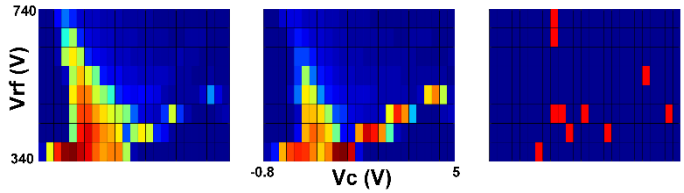


Figure 7. Example dispersion plot of muscle tissue (left), adipose tissue (center), and ten features selected in the FSFS process for the classification of the tissues (right). Vrf is the DMS separation voltage and Vc is the DMS compensation voltage.

4. Discussion

a) Findings and impact

Our results show that the latency, classification accuracy and reliability of our freehand-operated tissue analysis system are acceptable and pave way towards clinical validation of the technology. The study works as a proof-of-concept for previously unexplored application of rapid DMS-based tissue identification. However, some limitations that hinder the integration of the system to a clinical setting still remain to be resolved.

b) Limitations

As this study focused on describing the technical novelty of the improved surgical system, we cannot emphasize the clinical relevance of the identification results and their impact. However, the simplified binary setting works as a promising surrogate for a real surgical situation since the intended application of the system would be to differentiate malignant and benign tissues in a similar manner. In addition, muscle and adipose tissue are usually the most prominent types of healthy tissue surrounding tumors, which makes their identification relevant.

In the intended final surgical application, one of the most desired attributes of the system is minimal analysis time, so that the surgeon can perform the operation without disturbances in the workflow. The current minimum interval between the DMS measurements, approximately 9 seconds, should be improved in order to make the assessment of the resection

line easier. With this respect, our system is still inferior to the current MS-based solutions. With REIMS, the time between the start of the cut to displaying the result is reported to be less than two seconds [13]. The limiting factor for the measurement duration in our system is the ENVI-AMC DMS. As most DMS devices are not designed to perform rapid measurements, but rather to monitor ambient air quality, the measurements require a full sweep of voltage values and cannot be concentrated on some key areas of the spectrum. Currently, the only way to decrease the measurement time is to decrease the resolution or the spectrum window size, which will decrease the classification performance.

Even in the task of a rather simple classification between skeletal muscle and intermuscular fat, the system failed to achieve 100% accuracy. This presents a clear area for improvement regarding the diagnostic properties of the system. However, the main reason for the misclassification can be traced back to the variability on the external measurement conditions that affect the DMS output. Especially the changes in environmental humidity and temperature can cause differences between the spectra in terms of offsets, signal intensities, and even feature positions. This is apparent in the CNN results of the independent data set, which were at guess level. To account for the variability of the measured outputs over time, a calibration protocol with a known control substance should be created in future studies.

The reference annotation of the tissue specimens was conducted by macroscopic examination and it is possible that a subset of the muscle specimens contained significant amount of fat that may impair the results. However, the effect is likely small, and the risk of overoptimistic results is low.

In addition, due to a physical blockage in the airways of the DMS device that occurred between the measurements of the training data and the independent test set, the DMS core circuit boards were disassembled, cleaned, and reassembled before measuring the independent test set. The reassembly might have caused some change in the measurement output.

c) Analysis of the Results

Besides pointing out the significant negative effect of long-term output drift of the DMS device to classification accuracy, the CNN and LDA results showed that the performance of the classification model is largely dependent on the size and variability of the training set. Even with data that were not produced by freehand surgery, the classification performance for the independent data set increased with both models, when the size of the training set was increased. Furthermore, the CNN model was particularly affected by the added training data, since the classification accuracy increased from guess level prediction to 88%, i.e. the number of misclassifications was decreased by 76%. This is perhaps expected, since neural networks are highly dependent on the amount of training data and they are most effective in applications, where the size of the training set can be millions of samples [26]. However, these results present interesting research questions regarding the optimal data analysis of this type of DMS data that should be answered in future studies.

Another aspect that needs further exploration is the effect of carry-over signal to classification performance. While the integrated heating of the sample pre-processing unit and PID-controlled dilution systems were implemented and reduced the overall recovery time of the system, residual scent of previous measurements can accumulate and slowly release from the altogether 3-meter-long tubing before the pre-processing unit. The carry-over signal might explain part of the misclassifications in this study. However, in similar applications based on MS, the study groups have been able to get near perfect classification results with rapid measurements [14], [19], indicating that the possible carry-over signal can be distinguished from the relevant information of a new sample, at least in MS-based solutions.

The relevant information of the samples can be considered the key features that enable the differentiation based on the DMS spectra. However, some areas of the spectra do not yield any significant information regardless of the measured substance, but contain electronic noise derived from the DMS hardware. In other words, the spectra contain changes that are not related to the tissue type and thus including them as training data for the classifier can predispose the model to overfitting and reduce the generalizability. Much like in our previous study with breast cancer samples, the FSFS results revealed that the acquisition of the full DMS spectrum from each smoke sample is redundant in terms of classification accuracy [21]. In this study, by selecting only 5% (10/200) of the features, the classification accuracy increased by one decimal compared to the classification with the full spectrum. This means that if the DMS device would allow for selection of specific waveform voltage and compensation voltage values instead of full sweeps, the measurement time could be significantly shortened, while retaining similar, or even better, performance. This, in turn, would lead to better applicability of the system to the intended real time surgical use.

Since the aim of the system is qualitative assessment of tissue type and extensive libraries for DMS are not available, the direct identification of the measured molecules is not required. However, the features selected by the FSFS are indicative of the cellular content differences between muscle and fat. The water content between the tissues is different and this can be seen in the selected features in the rightmost selected features that are in line with the so-called reactant ion peak. The selected features on the left side of the spectra derive from differences in heavier molecules such as lipids and their degradation products.

5. Conclusion

We have shown that fast tissue identification during freehand diathermy cutting based on diathermy smoke analysis by DMS is possible. In addition to the previously used LDA classification, we employed a CNN model that has potential to improve with the addition of new data. In terms of diagnostic accuracy and overall performance, the DMS-based system is not yet capable of producing results that would compete with the MS-based systems, but its simplicity and robustness partly compensate its limitations. The difficulties associated with freehand smoke sampling have been mostly overcome with engineering solutions and the system will continue to improve with further research. In future studies, the performance of the system will be tested in a surgical setting with clinically relevant human tissues and the following results will determine the feasibility of the technology in clinical use.

References

[1] A. Bodilsen *et al.*, “The Influence of Repeat Surgery and Residual Disease on Recurrence After Breast-Conserving Surgery: A Danish Breast Cancer Cooperative Group Study,” *Ann. Surg. Oncol.*, vol. 22, no. 3, pp. 476–485, 2015.

[2] T. P. Ahern, H. Larsson, J. P. Garne, D. P. Cronin-Fenton, H. T. Sørensen, and T. L. Lash, “Trends in breast-conserving surgery in Denmark, 1982–2002,” *Eur. J. Epidemiol.*, vol. 23, no. 2, pp. 109–114, 2008.

[3] R. Jeevan *et al.*, “Reoperation rates after breast conserving surgery for breast cancer among women in England: Retrospective study of hospital episode statistics,” *BMJ*, vol. 345, no. 7869, pp. 1–9, 2012.

[4] S. Jendrian, K. Steffens, B. Schmalfeldt, E. Laakmann, C. Bergelt, and I. Witzel, "Quality of life in patients with recurrent breast cancer after second breast-conserving therapy in comparison with mastectomy: the German experience," *Breast Cancer Res. Treat.*, vol. 163, no. 3, pp. 517–526, 2017.

[5] Y. Grant *et al.*, "Patient-level costs in margin re-excision for breast-conserving surgery," *BJS*, vol. 106, no. 4, pp. 384–394, Mar. 2019.

[6] R. E. Pataky and C. R. Baliski, "Reoperation costs in attempted breast-conserving surgery: A decision analysis," *Curr. Oncol.*, vol. 23, no. 5, pp. 314–321, 2016.

[7] J. C. Boughey *et al.*, "Impact of analysis of frozen-section margin on reoperation rates in women undergoing lumpectomy for breast cancer: Evaluation of the National Surgical Quality Improvement Program data," *Surg. (United States)*, vol. 156, no. 1, pp. 190–197, 2014.

[8] J. M. Jorns *et al.*, "Intraoperative frozen section analysis of margins in breast conserving surgery significantly decreases reoperative rates: One-year experience at an ambulatory surgical center," *Am. J. Clin. Pathol.*, vol. 138, no. 5, pp. 657–669, 2012.

[9] E. R. St John *et al.*, "Diagnostic accuracy of intraoperative techniques for margin assessment in breast cancer surgery a meta-analysis," *Ann. Surg.*, vol. 265, no. 2, pp. 300–310, 2017.

[10] S. J. Erickson-Bhatt, R. M. Nolan, N. D. Shemonski, S. G. Adie, J. Putney, and D. Darga, "Real-time imaging of the resection bed using a handheld probe to reduce incidence of microscopic positive margins in cancer surgery," *Cancer Res*, vol. 75, no. 18, pp. 3706–3712, 2015.

[11] S. A. Boppart *et al.*, "Label-free optical imaging technologies for rapid translation and use during intraoperative surgical and tumor margin assessment," *J. Biomed. Opt.*, vol. 23, no. 2, pp. 1–10, 2018.

[12] F. Schnabel *et al.*, "A randomized prospective study of lumpectomy margin assessment with use of marginprobe in patients with nonpalpable breast malignancies," *Ann. Surg. Oncol.*, vol. 21, no. 5, pp. 1589–1595, 2014.

[13] E. R. St John *et al.*, "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.*, vol. 19, no. 59, pp. 1–14, 2017.

[14] D. L. Phelps *et al.*, "The surgical intelligent knife distinguishes normal, borderline and malignant gynaecological tissues using rapid evaporative ionisation mass spectrometry (REIMS)," *Br. J. Cancer*, vol. 118, no. 10, pp. 1349–1358, 2018.

[15] K. O. Alfarouk and C. T. S. and L. Schwartz, "The Warburg Effect and the Hallmarks of Cancer," *Anti-Cancer Agents in Medicinal Chemistry*, vol. 17, no. 2. pp. 164–170, 2017.

[16] B. Fatou *et al.*, "In vivo Real-Time Mass Spectrometry for Guided Surgery Application," *Sci. Rep.*, vol. 6, no. 25919, pp. 1–14, May 2016.

[17] M. Woolman *et al.*, "Rapid determination of medulloblastoma subgroup affiliation with mass spectrometry using a handheld picosecond infrared laser desorption probe," *Chem. Sci.*, vol. 8, no. 9, pp. 6508–6519, 2017.

[18] J. Zhang *et al.*, "Nondestructive tissue analysis for ex vivo and in vivo cancer diagnosis using a handheld mass spectrometry system," *Sci. Transl. Med.*, vol. 9, no. 406, pp. 1–11, Sep. 2017.

[19] P. Saudemont *et al.*, "Real-Time Molecular Diagnosis of Tumors Using Water-Assisted Laser Desorption/Ionization Mass Spectrometry Technology," *Cancer Cell*, vol. 34, no. 5, pp. 840–851, 2018.

[20] A. Kontunen, M. Karjalainen, J. Lekkala, A. Roine, and N. Oksala, "Tissue Identification in a Porcine Model by

Differential Ion Mobility Spectrometry Analysis of Surgical Smoke," *Ann. Biomed. Eng.*, vol. 46, no. 8, pp. 1091–1100, 2018.

[21] M. Sutinen *et al.*, "Identification of breast tumors from diathermy smoke by differential ion mobility spectrometry," *Eur. J. Surg. Oncol.*, vol. 45, no. 2, pp. 141–146, Feb. 2019.

[22] I. Haapala *et al.*, "Identifying brain tumors by differential mobility spectrometry analysis of diathermy smoke," *J. Neurosurg. JNS*, vol. 1, no. AOP, pp. 1–7, 2019.

[23] M. Karjalainen *et al.*, "The characterization of surgical smoke from various tissues and its implications for occupational safety," *PLoS One*, vol. 13, no. 4, pp. 1–13, Apr. 2018.

[24] O. Anttalainen, J. Puton, K. Peräkorpi, E. Budzyńska, G. Eiceman, and M. Sillanpää, "Differential mobility spectrometers with tuneable separation voltage – Theoretical models and experimental findings," *TrAC Trends Anal. Chem.*, vol. 105, no. 8, pp. 413–423, 2018.

[25] S. Wang, J. Tang, and H. Liu, "Feature Selection BT," in *Encyclopedia of Machine Learning and Data Mining*, C. Sammut and G. I. Webb, Eds. Boston, MA: Springer US, 2017, pp. 503–511.

[26] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, no. 5, pp. 354–377, 2018.

PUBLICATION V

Tissue Identification from Surgical Smoke by Differential Mobility Spectrometry: an in vivo study

Anton Kontunen*, Ulla Karhunen-Enckell*, Markus Karjalainen, Anna Anttalainen, Pekka Kumpulainen, Leena Pitkänen, Osmo Anttalainen, Antti Vehkaoja, Niku Oksala, and Antti Roine

*Equal contribution

IEEE Access, 9, 168355-168367, 2021.

DOI: https://doi.org/10.1109/ACCESS.2021.3136719

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 26, 2021, accepted December 12, 2021, date of publication December 20, 2021,
date of current version December 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3136719

Tissue Identification From Surgical Smoke by Differential Mobility Spectrometry: An in Vivo Study

ANTON KONTUNEN[1,2], ULLA KARHUNEN-ENCKELL[1,3], MARKUS KARJALAINEN[1,2], ANNA ANTTALAINEN[2,4], PEKKA KUMPULAINEN[2], LEENA PITKÄNEN[3], OSMO ANTTALAINEN[2], ANTTI VEHKAOJA[1,5], (Member, IEEE), NIKU OKSALA[1,2,5,6], AND ANTTI ROINE[2]

[1]Faculty of Medicine and Health Technology, Tampere University, 33100 Tampere, Finland
[2]Olfactomics Oy, 33720 Tampere, Finland
[3]Department of Surgery, Tampere University Hospital, 33520 Tampere, Finland
[4]Medaffcon Oy, 02130 Espoo, Finland
[5]Finnish Cardiovascular Research Center, 33100 Tampere, Finland
[6]Vascular Centre, Tampere University Hospital, 33520 Tampere, Finland

Corresponding author: Anton Kontunen (anton.kontunen@tuni.fi)

The work of Anton Kontunen was supported in part by the Doctoral School of Tampere University, in part by the Finnish Foundation for Technology Promotion under Grant 7671, and in part by the Emil Aaltonen Foundation under Grant 210073. The work of Markus Karjalainen was supported by the Finnish Cultural Foundation through the Pirkanmaa Regional Fund. The work of Niku Oksala was supported in part by the Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital under Grant 9s045, Grant 9T044, Grant 9U042, Grant 150618, Grant 9V044, Grant 9X040, Grant 9AA057, Grant 9AB052, and Grant MK301; in part by the Academy of Finland under Decision 292477 (to strengthen university research profiles); and in part by the Tampereen Tuberkuloosisäätiö (Tampere Tuberculosis Foundation).

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Tampere University Hospital under Approval No. R17096, and performed in line with the World Medical Association's Declaration of Helsinki.

ABSTRACT The increasing number of breast cancer survivors and their longevity has emphasized the importance of esthetic and functional outcomes of cancer surgery and increased pressure for the surgical treatment to achieve negative margins with minimal removal of healthy tissue. Surgical smoke has been successfully utilized in tissue identification in laboratory conditions by using a system based on differential mobility spectrometry (DMS) that could provide a seamless margin assessment method. In this study, a DMS-based tissue analysis system was used intraoperatively in 20 breast cancer surgeries to assess its feasibility in tissue identification. The effect of the system on complications and duration of surgeries was also studied. The surgeries were recorded with a head-worn camera system for visual annotation of the operated tissue types to enable classification of the measurement files by supervised learning. There were statistically significant differences among the DMS spectra of the tissue types. The classification of four tissue types (skin, fat, glandular tissue, and connective tissue) yielded a cross-validated accuracy of 44% and exhibited high variation between surgeries. The low accuracies can be attributed to the limitations and uncertainty of the visual annotation, high-within class variance due to the heterogeneity of tissues as well as environmental conditions, and delays of the real-time analysis of the smoke samples. Differences between tissues encountered in breast surgery were identified and the technology can be implemented in surgery workflow. However, in its current state, the DMS-based system is not yet applicable to a clinical setting to aid in margin assessment.

INDEX TERMS Biomedical engineering, biomedical measurement, breast cancer, differential mobility spectrometry, supervised learning, surgical instruments, surgical margin, surgical smoke.

The associate editor coordinating the review of this manuscript and approving it for publication was György Eigner.

I. INTRODUCTION

Breast cancer is the most common cancer affecting more than two million women worldwide annually [1]. The prognosis of early-stage breast cancer is good – more

than 90% of the patients are alive five years after diagnosis [2], [3] Thus, besides oncological outcome, esthetic and functional outcomes are becoming increasingly important for patients as a factor of quality of life and overall health [4], [5].

As tumors are found earlier and smaller, and larger tumors may be operated utilizing oncoplastic methods, more patients are likely to receive breast conserving therapy, a combination of breast conserving surgery (BCS) and whole breast irradiation to eradicate any microscopic residual disease. The majority of breast cancers are treated with breast conserving therapy in Europe [6] and the United States [7].

The aim of BCS is to remove the tumor with histologically negative margins. According to current guidelines, a negative margin is defined as no ink on tumor for invasive carcinoma and 2 mm histological margin for ductal carcinoma in situ (intraductal carcinoma, DCIS) [8]–[11]. Although acceptable margins are narrow, the tumor is resected with larger margins, due to unevenness of the tumor borders and inability to assess borders intraoperatively. Positive histological margin increases the risk of local recurrence [11] and reoperation is recommended to obtain negative margins. The average reoperation rate is approximately 20% but it varies widely from less than 10% [12] to more than 60% [13] between surgeons, facilities, and cancer types [14]–[18]. Reoperations may worsen the prognosis by delaying adjuvant therapy [19], cause psychological stress, and impair the cosmetic outcome of the treatment [20]. Reoperations are also associated with higher incidence of post-operative wound complications [21] and increased economic burden [22]. On the other hand, patients with smaller excision volumes have improved cosmetic outcomes compared to larger excision volumes [20], [23].

The resection volumes and margins can be optimized by intraoperative margin assessment. The surgical specimen can be assessed by x-ray (specimen radiography) or ultrasound to ensure that the radiologically visible tumor has been removed with sufficient radiological margins. They enable the assessment of radiologically visible borders, but not microscopic borders. Microscopic assessment is traditionally carried out by frozen section analysis or imprint cytology of the resection margins. Both techniques are time-consuming and resource-intensive, and their use is limited [24]. A solution based on Radiofrequency spectroscopy has been approved by the Food and Drug Administration to provide intraoperative evaluation of the tissue at the edges of excised breast tissue. The device measures the local electrical properties of breast tissue, which differ between normal and malignant tissue [25] and provides a positive or negative reading for each measurement taken [26]. It has been shown to reduce reoperation rates [27] but has not reached wide clinical adoption [13].

Experimental methods based on optical imaging and mass spectrometry (MS), have shown promise in terms of applicability for intraoperative margin detection. Among the optical methods, optical coherence tomography and photoacoustic tomography have achieved sensitivities of over 90% [28], [29]. Their shortcomings are the expertise needed for image interpretation and reliance on ex vivo analysis of the specimen. MS analyzes the molecular content of the specimen. Several MS-based techniques have consistently exhibited classification accuracies of 90% in tissue identification and detection of different cancer types [30]–[33]. Of these methods, Rapid evaporative ionization mass spectrometry (REIMS), which is based on MS analysis of surgical smoke, has been the most extensively studied, and it has been proven to be capable of real time analysis in in vivo studies [34]. The cost, complexity and large physical size limit the clinical applicability of MS-based methods. Differential mobility spectrometry (DMS) is a technology that separates gaseous substances at a molecular level by ionizing the sample at atmospheric pressure, after which the sample ions are separated and measured based on their mobility characteristics in an asymmetrical high voltage electric field [35]. Due to its freedom from the requirement of a vacuum, and less complex design, DMS sensors are more affordable and smaller than MS instruments, which improves their adaptability to a clinical setting.

In previous studies, DMS-based tissue identification from surgical smoke has been tested in laboratory conditions, where the classification with several porcine tissues has yielded accuracy results of over 90% [36]–[38]. In a laboratory study on human breast cancer identification, the DMS-based method achieved a classification accuracy of 87 % between benign and malignant tissues [39]. While these results are promising, the applicability and performance of the technology has not yet been demonstrated in vivo in a clinical setting. The effect of variation in environmental factors and sampling in clinical use on the method remains unknown. The establishment of sufficient dataset from positive margins in vivo would require a significant number of patients as intentional creation of positive margins is not ethically feasible in human studies and the positive margin rate in our institution is around 10%. Additionally, the annotation of tissues in intraoperative use is not trivial and requires innovative approaches. For these reasons, with a pilot study of 20 patients, we concentrated on the feasibility of use of the introduced technology and its abilities to identify benign tissues, rather than its margin assessment performance. We also demonstrate a novel, minimally intrusive, intraoperative tissue annotation method based on video footage captured from the point of view of the operating surgeon.

II. METHODS

A. PATIENTS AND CLINICAL DATA

This was a prospective single-arm first-in-human single-center study performed between 9th of October and 26th of November 2019 at Tampere University Hospital, Finland. In this study, the operating surgeon was blinded from the measurement results and the measurements were not used to assess the margins or guide the operation. Ethical approval was obtained from the local Ethics Committee of Tampere University Hospital (code R17096). The study was conducted

in accordance with the World Medical Association's Declaration of Helsinki. Informed consent was obtained from all patients in written. Inclusion criteria included patients over 18 years recently diagnosed with any histological type of invasive breast cancer or DCIS or atypical ductal hyperplasia and who were eligible for BCS. Patients with impalpable lesions underwent ultrasound- or mammography-guided wire localization preoperatively. All breast and axillary operations were performed by two experienced breast-cancer-dedicated plastic surgeons. Operations were carried out following the national guideline [40], which is in line with international European and North American guidelines [8], [10], [11], [41]. The removed breast tissue was assessed grossly in case of palpable lesions and via specimen radiography if the tumor was localized with a wire. Additional breast tissue was removed if a positive margin was suspected.

Patient data was collected from electronic health records and operation times obtained from operation room management system. Histological data was gathered from structured histopathology report. Tumor volume was calculated using the diameter of the tumor (in cm) as mentioned in the pathology report and assuming spherical shape. Total resection volume (TRV) was calculated using three dimensions of the surgical specimen and assuming ellipsoidal shape. The optimal resection volume (ORV) was defined as the spherical volume of the tumor itself with an added 1.0 cm margin of healthy breast tissue. The method was adapted from the study by Krekel *et al.* [18]. Oncoplastic reduction mammoplasties were discarded from volumetric calculations because excessive amount of breast tissue is excised due to operation technique rather than to remove the tumor with adequate margins. The time from the first incision to closure and total operation room time of the surgeries were compared to institutional averages by one sample t-test.

B. MEASUREMENT SYSTEM

Automatic tissue analysis system (ATAS), previously described in Kontunen *et al.* 2021 [38], was used in the study. The function of ATAS is based on a surgical smoke preprocessing unit and a DMS sensor. The system can, with minor modifications, be attached to any commercially available diathermy units and smoke evacuation devices. In short, the operation principle of ATAS is as follows: 1) Surgeon operates tissue with a diathermy instrument. 2) Induced current from the dispersive electrode of the diathermy system is measured by an encased induction coil. If the induced current exceeds a pre-determined threshold, the system interprets that tissue has been cut and triggers a DMS measurement. 3) A small fraction of the surgical smoke sample that is evacuated from the surgical area is taken into the pre-processing unit where it is diluted and filtered by an electric filter to remove contaminating particulate matter. 4) The filtered sample is measured by the DMS sensor (ENVI-AMC, Environics Oy, Finland) approximately five seconds after the trigger signal has been received. The measurement data are stored to the local database of the system, after which the process (starting from step 1) can repeat for a subsequent measurement. In this study, the operating surgeon was blinded from the measurement results and the measurements were not used to assess the margins or guide the operation.

The duration of the DMS measurement is approximately 5 seconds, which means that together with the start delay of 5 seconds, the minimum time for one measurement cycle is 10 seconds. As its output, the system produces a measurement file that contains the DMS measurement data, the diathermy current measurement data, and current measurement data from the electric filter. The system and its simplified operation principle are depicted in Fig. 1. For a more in-depth description and schematic representations of the measurement system, the reader is referred to Kontunen *et al.* [38]. The diathermy power unit that was used alongside the system was a *Berchtold Elektrotom 530 Electrosurgical Unit* (Stryker Corp, USA) and the surgical smoke evacuator was a *SafeAir® Smoke Evacuator compact* (Stryker Corp, USA) that was operated at a power setting of 7/10 and in continuous evacuation mode.

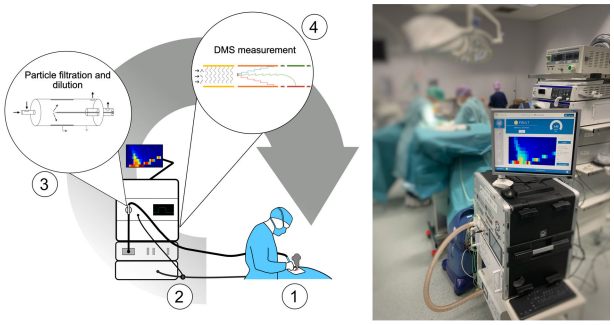


FIGURE 1. Simplified operation principle of the measurement system (left) and the measurement system in an operation room (right). 1) Surgical smoke is produced by the diathermy instrument. 2) Induced current from the dispersive electrode triggers the measurement. 3) Surgical smoke sample is taken into the pre-processing unit for particle filtering and dilution. 4) The filtered sample is measured by the DMS sensor, and the result is stored to internal memory. The duration of one measurement loop is approximately 10 seconds.

C. ANNOTATION SYSTEM AND STATISTICAL TOOLS

The result of the smoke sample measurement was not interpreted in real time in this study, since the study functioned as a pilot for *in vivo* DMS-based tissue identification, and thus a pre-trained model for tissue classification was not available. Instead, the data produced during the surgeries was annotated and classified post-operatively based on video footage of the operation. Each surgery was recorded by a head-mounted camera (Pupil Core, Pupil Labs GmbH, Germany) that was worn by the operating surgeon [42]. The video footage recorded by the camera was stored locally to a dedicated mobile device (Motorola XT1929-8 Moto Z3 Play, Motorola, USA) during the surgery, from which the footage was transferred to an encrypted hard disc drive for storage and later data analysis.

The statistical analysis was done in R software environment [43], integrated development environment

RStudio [44], and Matlab (version R2019a, MathWorks, USA). Before tissue annotation, the raw video data was also processed with Matlab. The video footage was synchronized with the measurement data by creating an individual video clip for each measurement based on the recorded time labels. In addition, each clip was also overlayed with the graphs and DMS spectrum of the corresponding measurement to aid in the timing for the annotation of the samples. An example still image of a video clip that was used in tissue annotation is presented in Fig. 2.

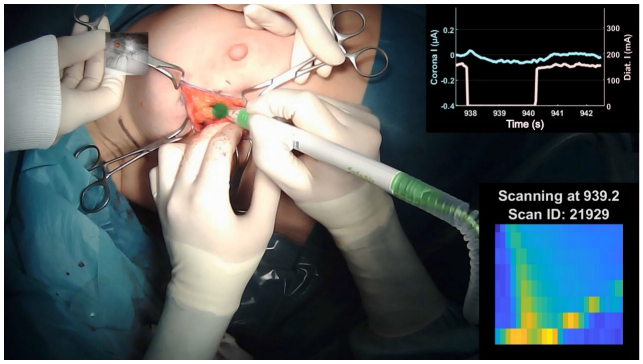


FIGURE 2. Still image of a video file that was used in tissue annotation. The DMS spectrum of the measurement is overlayed in the bottom right corner and the current of the electric corona filter and the measured induction current from the dispersive diathermy electrode can be seen in the upper right corner. The round dot in the incision area is from the gaze tracking feature of the Pupil core camera.

D. ANNOTATION WORKFLOW

In total, the number of individual video clips was 1131. Due to the high amount of annotatable data, the totality of the video material was annotated by only one medical expert. To investigate the potential subjectivity of only one observer, the inter-rater agreement in the video-based annotation was studied with a subset of the measurement data.

The annotation based on the video footage was initially tested by observing two surgeons as they viewed and annotated 30 samples of a randomly selected operation. Without specific instructions, the variation between the annotators was high in terms of terminology, assigned class and determination of sufficient sample. Thus, a protocol for the video annotation for the full data set was made.

According to the protocol, the viewer should assign the type of the measured tissue, when the DMS measurement was initiated, i.e., approximately five seconds after the trigger signal has been received. In an annotatable video clip, this was indicated by the appearance of the spectrum to the lower right corner (Fig. 2). The possible assigned tissue classes were determined as: skin, fat, glandular tissue, connective tissue, muscle, blood, and empty (i.e., no cutting occurs during DMS measurement). In addition, the protocol stated that the viewer should evaluate the sufficiency of the smoke sample and add notes regarding possible irregularities. The visual sufficiency estimate was included in the protocol as a possible exclusion criterium for final analysis. Table 1 shows an example output of the annotation process for four measurements.

TABLE 1. Example annotations.

Annotator: Surgeon 1			
Video clip	Tissue	Sample quantity	Notes
Part 23	Empty	Insufficient	No cutting during DMS measurement
Part 24	Skin	Sufficient	
Part 25	Empty	Insufficient	The diathermy was activated accidentally in its holding bag
Part 26	Glandular	Sufficient	

The efficiency of the annotation protocol in terms of inter-rater agreement was estimated by the Fleiss' kappa metric [45]. In practice, this means that a randomly selected statistically sufficient portion of the measurement footage was annotated by three individuals, after which the Fleiss' kappa was calculated for the annotation matrix to see the rate of agreement. The power calculations were done by utilizing the R package k*appaSize* by Rotondi [46]. The sample size for the inter-rater agreement was based on a power calculation with the following parameters: the null hypothesis (kappa0) for the kappa test was set to 0.01, the alternative hypothesis (kappa1) to 0.2, the type I error rate (alpha) to 0.001, and the desired level of statistical power (power) to 0.95. The anticipated prevalence of different classes (props) was estimated based on already completed annotations by one observer.

E. DATA ANALYSIS

The full annotation data from one observer was utilized in further statistical analysis and tissue classification. However, additional data curation was deemed necessary due to the highly variable nature of the diathermy activations during surgery and failed measurements. In the first two surgeries, the surgical evacuator was operated with the maximum power, but the pressure ejector system in the sample pre-processing unit was not optimized to overcome the suction, i.e., the entirety of the smoke went to the surgical evacuator. In operations 13 and 14, the video data was not saved due to a malfunction in the process of saving the data to the memory card of the mobile device from the Pupil camera system. The exclusion criteria are presented in Fig. 3. After all exclusion steps, the number of measurement files was 611 (fat, N = 395; glandular tissue, N = 129; skin, N = 52; connective tissue, N = 35).

The final DMS measurement files were classified with a regularized linear discriminant analysis (LDA). LDA is a relatively simple supervised method that tries to

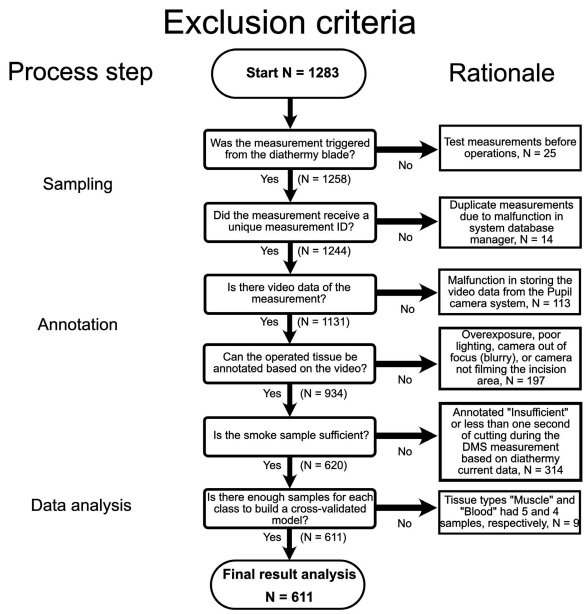


FIGURE 3. The exclusion criteria for the DMS measurement data.

maximize class separation based on a linear projection of the feature space [47]. LDA assumes equal covariance between the classes and the class of a sample is determined based on its distance from the class mean after the linear projection. LDA has previously been used both in MS and DMS-based tissue classification [34], [39]. In this study, the classification was done based on the DMS spectra that consisted of the measured values of positive ions. Each spectrum was measured with the DMS compensation field voltages of −0.8 V to 5 V, in 25 steps and separation field voltages of 340 V to 740 V, in 8 steps, resulting in 200 values for each DMS measurement. The classification performance was analyzed with leave-one-surgery-out cross-validation to alleviate overfitting. However, due to the unbalanced ratio between the tissue types and variation between surgeries, each surgery was also classified individually using leave-one-sample-out cross-validation.

To further analyze the differences between the DMS spectra of the tissue types, the distributions of the dispersion plot values were subjected to the Kolmogorov-Smirnov test to identify features that are statistically different among the classes [48]. The statistical significance was determined at a significance level of 0.05 and the p-values were Bonferroni-corrected by the number of dimensions (200). This means that a p-value of 0.00025 was considered statistically significant.

III. RESULTS

A total of 20 women were operated. A summary of demographic data and clinical characteristics are depicted in Table 2. Four patients (20%) underwent lumpectomy, ten patients (50%) level 1 oncoplastic breast conserving surgery, one patient (5%) oncoplastic breast conserving surgery combined with reduction mammoplasty of the healthy breast, and five patients (25%) oncoplastic reduction mammoplasty combined with reduction mammoplasty of the healthy breast. Surgeon 1 operated 13 (65%) patients and surgeon 2 seven (35%) patients. Average operating time from skin incision to skin closure and total operation room time were similar to the institutional average (Table 2) as determined by the one sample t-test, which produced p-values that indicated no statistically significant difference between the means.

On histopathological analysis, all but one patient had sufficient histological margins both from invasive ductal carcinoma (IDC) and DCIS. One patient, diagnosed preoperatively with 8 mm grade 1 pure DCIS, had DCIS grade 1 sized 11 mm with positive lateral margin on histopathological analysis. The patient had re-resection and final pathological analysis revealed 6 mm more of DCIS grade 1 but the margins were sufficient. One patient had a diagnosis of 2 mm pleomorphic lobular carcinoma in situ (LCIS) on final histopathological analysis, although preoperative diagnosis had been DCIS grade 3, and smallest lateral 1 mm margin was accepted. Therefore, reoperation rate was 5%. Excluding oncoplastic reduction mammoplasties the average lateral margin on histopathological analysis was 16.4 mm from IDC and 11.5 mm from DCIS. Anterior and posterior margins were sufficient in all cases. The extent of lateral margin widths is further depicted in Table 3. Excluding oncoplastic reduction mammoplasties, the amount of tissue removed was 49.3 cm^3 on average, when 24.2 cm^3 would have been theoretically optimal if 1 cm macroscopic margins were used. The TRV:ORV-ratio was thus, on average, 2.0.

Closest lateral tumor margin from both invasive and intraductal carcinoma (all tumors), from invasive ductal carcinoma (IDC) and from pure ductal carcinoma in situ (DCIS) or DCIS component of invasive carcinoma. DCIS includes one patient with pleomorphic lobular carcinoma in situ. Patients receiving oncoplastic reduction mammoplasty were excluded.

A. COMPLICATIONS

One patient (5%) suffered wound infection and dehiscence postoperatively. The patient was treated conservatively in outpatient setting and received oral antibiotics. One patient (5%) had a small hematoma that required one postoperative puncture. One patient out of 18 (5.6%), who underwent sentinel lymph node biopsy, suffered lymphedema of the ipsilateral arm. The patient received both physical therapy and compression garment.

B. INTER-OBSERVER AGREEMENT

From the total of 1131 annotatable video files, the inter-rater agreement assessment was done with a subset of 72 files. For these files, the three observers annotated the samples to five classes as instructed. The inter-rater agreement results based on the Fleiss' kappa metric are presented in Table 4.

TABLE 2. Summary of patient demographics and clinical characteristics.

		N	Mean	Range	Percentage
General	**Age (years)**	-	64.2	50–80	-
	Body mass index, kg/m^2	-	28.1	20.8–38.3	-
Tumor size on imaging (mm)	**Mammogram**	-	13.6	6–25	-
	Ultrasound	-	14.2	6–40	-
	MRI	-	23.7	11–50	-
Tumor localization	**Palpation**	3	-	-	15%
	Wire-guided localization	17	-	-	85%
Final diagnosis	**IDC**	16	-	-	80%
	DCIS	3	-	-	15%
	LCIS	1	-	-	5%
Tumor size (greatest dimension)	**< 10 mm**	4	-	-	20%
	10–20 mm	12	-	-	60%
	> 20 mm	4	-	-	20%
Follow-up	**Postoperative infection**	1	-	-	5%
	Reoperations	1	-	-	5%
Duration of lumpectomy (h:min)	**Incision–closure**	-	0:53 (Institutional average 1:04)	-	-
	One sample t-test p-value	-	0.20	-	-
	Operation room time	-	1:44 (Institutional average 1:47)	-	-
	One sample t-test p-value	-	0.66	-	-
Duration of oncoplastic BCS (h:min)	**Incision–closure**	-	1:50 (Institutional average 1:42)	-	-
	One sample t-test p-value		0.60	-	-
	Operation room time	-	2:44 (Institutional average 2:31)	-	-
	One sample t-test p-value	-	0.42	-	-

Abbreviations: MRI (magnetic resonance imaging), IDC (invasive ductal carcinoma), DCIS (ductal carcinoma in situ), LCIS (lobular carcinoma in situ).

C. CLASSIFICATION

The leave-one-patient-out cross-validated LDA classification of the accepted dataset produced a mean classification accuracy of 44.3% for four tissue types (skin, fat, glandular, and connective tissue). The results of the leave-one-sample-out cross-validated LDA classification of each surgery are presented in Table 5.

Pair-wise comparison between tissue classes revealed statistically significant differences between fat and the other tissue types, and between skin and glandular tissue.

TABLE 3. Closest lateral tumor margins.

	All tumors		IDC		DCIS	
Positive	1	6.7%	0	0%	1	14.3%
No ink on margin–1.9 mm	1	6.7%	0	0%	1	14.3%
2–4.9 mm	3	20.0%	2	18.2%	1	14.3%
5–9.9 mm	4	26.7%	4	36.4%	1	14.3%
10–19.9 mm	4	26.7%	3	27.3%	3	42.9%
> 20 mm	2	13.3%	2	18.2%	0	0%
Total	15		11		7	

Closest lateral tumor margin from both invasive and intraductal carcinoma (all tumors), from invasive ductal carcinoma (IDC) and from pure ductal carcinoma in situ (DCIS) or DCIS component of invasive carcinoma. DCIS includes one patient with pleomorphic lobular carcinoma in situ. Patients receiving oncoplastic reduction mammoplasty were excluded.

TABLE 4. Fleiss' kappa for the inter-rater agreement of three observers and 72 samples.

	Skin	Fat	Glandular	Connective Tissue	Empty
Kappa for class	0.7126	0.3649	0.3561	0.3709	0.5350
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Fleiss' kappa	0.4198				
Confidence interval	0.3982–0.4415				
Agreement*	Moderate				

*Agreement based on Landis, J. R. and Koch, G. G. (1977) [65]

However, there were no statistically significant difference in the features between glandular and connective tissue, or skin and connective tissue. The differing dispersion plot features along with an example DMS dispersion plot are highlighted in Fig. 4.

IV. DISCUSSION

The study showed that DMS can be implemented into the surgical workflow and demonstrated differences between signals from different tissues. The classification accuracy in this study did not reach the level of laboratory-based ex vivo studies [34], [38]. This underlines the challenges of macroscopic annotation of tissues and the importance of rapid measurement speed to obtain reliable results.

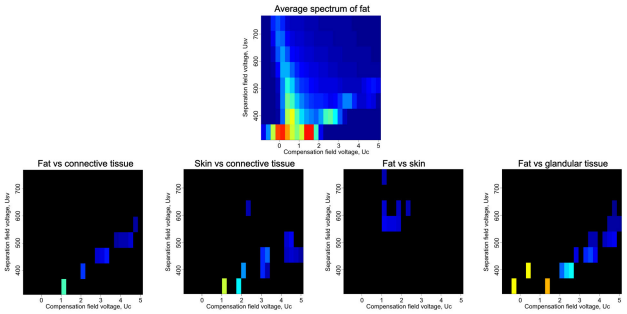


FIGURE 4. An average DMS spectrum of fat alongside features of the spectra that have a statistically significant difference between the tissue classes.

We identified statistically significant different features between the spectrums of most tissue types. The most relevant is the difference between glandular and fat tissue that are met concurrently in areas that are also prone to positive margins. Despite differences in spectra, there was wide variation in classification accuracy, ranging from 37% to 100%. The overall classification result with leave-one-surgery-out cross-validation was 44% with four tissue types. The leave-one-sample-out cross-validated results for each individual surgery showed that there is high variance in the classification accuracy between surgeries. This is largely explained by the varying number of accepted measurements between the surgeries and the annotated tissue classes. The highest accuracy was naturally acquired in surgeries, where the number of classes was two or three, since the classification problem was simplified. Studies on MS-based methods have approached the classification problem differently and the *in vivo* results are often not reported in terms of diagnostic accuracy, but rather as comparisons and statements that *in vivo* MS spectra were successfully acquired and that they resemble the *ex vivo* measurements [32], [34], [49]. However, in a recent REIMS study, a diagnostic accuracy of 90% was reported for binary *in vivo* classification of diseased and non-diseased rectal tissues in transanal minimally invasive surgery [33]. Thus, even though the study material and setting are different, it is apparent that the diagnostic performance of DMS in its current form does not match that of MS devices. However, the DMS sensor that was used in this study, is a prototype device that was initially designed for longer term monitoring of volatile organic and inorganic compounds, rather than rapid measurements of surgical smoke. By optimizing the sensor hardware for the specific medical application, better results could likely be achieved.

As the DMS-based system is ultimately intended for intraoperative margin assessment, it makes sense to compare it to techniques already available. Specimen radiography is widely used for documenting the removal of the targeted lesion but is conducted ex vivo and cannot be utilized continuously during the surgery. It also does not clearly improve the rates of reoperation for positive margins [50]. Ultrasound improves situational awareness of the surgeon and significantly reduces margin involvement and excision

TABLE 5. The leave-one-sample-out classification results for each surgery.

Surgery	Classification accuracy	N	Number of classes	Tissue classes (N per class)
1	No data	0	0	No data
2	No data	0	0	No data
3	65.4%	52	4	Fat (26), glandular (20), skin (3), connective tissue (3)
4	67.5%	40	4	Fat (22), glandular (8), skin (2), connective tissue (8)
5	65.2%	46*	3	Fat (34), skin (6), connective tissue (6)
6	100.0%	14*	2	Fat (3), glandular (11)
7	68.8%	64	4	Fat (45), glandular (12), skin (5), connective tissue (2)
8	80.8%	26*	3	Fat (9), glandular (14), skin (3)
9	66.7%	39	4	Fat (25), glandular (5), skin (6), connective tissue (3)
10	63.0%	73	4	Fat (61), glandular (3), skin (4), connective tissue (5)
11	60.5%	43	4	Fat (29), glandular (4), skin (5), connective tissue (5)
12	91.7%	36	3	Fat (25), glandular (9), skin (2)
13	No data	0	0	No data
14	No data	0	0	No data
15	75.0%	16	3	Fat (7), glandular (7), skin (2)
16	36.8%	19	4	Fat (6), glandular (9), skin (2), connective tissue (2)
17	87.5%	48	3	Fat (39), glandular (6), skin (3)
18	70.6%	17	3	Fat (12), glandular (3), skin (2)
19	88.6%	44	3	Fat (33), glandular (9), skin (2)
20	48.4%	31	3	Fat (19), glandular (8), skin (4)

*Operations 5, 6, and 8 had only one instance of glandular, skin, and connective tissue, respectively. These classes were omitted from the analysis, since leave-one-sample-out cross-validation requires at least two samples per class.

volumes in palpable and non-palpable tumors [51]–[53]. However, only half of the nonpalpable lesions can be visualized by ultrasound [54]. The use of intraoperative pathology requires pathology expertise intraoperatively and does not provide continuous feedback [50], [55]. Radiofrequency spectroscopy is fast (appr. 5–7 minutes), can be used by the surgeon and has achieved more than 50% reduction of reoperation rates for both invasive and intraductal carcinoma [26], [56]. Its shortcoming is the need for an additional probe during resection and a disruption of workflow to examine the resected tissue during the operation.

There remains a need for a reliable, fast, and cost-effective method for intraoperative assessment of surgical margins to reduce the rate of reoperations, excised breast volumes and mastectomy rates. An ideal device is coupled to the resection tool so that there is no interruption to standard workflow. It should detect cancer cells in real time guiding the resection *in vivo* and offer the surgeon the possibility to alter tissue excision. Moreover, it should be cost-effective and affordable globally. DMS-based system functioned reliably throughout the twenty surgeries. The malfunctions faced were all related to human errors or annotation system. The overall response of the operation room staff towards the system was positive. However, the noise of the pump system in the sample preprocessing unit was criticized by persons who were in close vicinity of the device during the operations. The system was placed near the patients' legs next to the surgical evacuator and diathermy unit, and while the size in its current state was

larger than the other instruments, the placement, preparation for use, and moving of the system between surgeries did not cause significant delays compared to normal operations.

The reoperation rate in our study was relatively low at 5%, but is in line with studies from another Finnish breast surgery unit [12], [57]. A study by Krekel et. al showed that TRV in BCS was median 2.5 times higher than would have theoretically been necessary to achieve sufficient margins [18]. In our series of lumpectomies and level 1 oncoplastic procedures the TRV:ORV ratio was 2.0, suggesting that low reoperation rate may at least partly be due to aggressive resections. The volume of excised tissue and the esthetic outcome of the operation are inversely proportional and excision volumes exceeding 50–85 cm^3 anticipate a cosmetic failure of BCS [23], [58]–[60]. On average, the resection volumes in this study remained under this threshold. Rate of surgical site infections and wound complications after breast cancer surgery vary from less than 5% to more than 30% depending on timing and definition [21], [61], [62]. In our series, the rate of infection or wound complications (5%), was similar to other studies [63], [64].

The reasons for the relatively low tissue classification results can be partially attributed to annotation. Firstly, the visual annotation of tissues based on video footage does not provide as definitive ground truth of the tissue class as histology does. There was a delay of 5 seconds between the reception of the trigger signal from the diathermy knife and the start of the DMS measurement, and the measurement itself took 5 seconds, meaning that the recorded signal represents the average of tissues operated in the period, which in some cases included more than one tissue. We found moderate agreement between the three annotators [65]. The inter-rater agreement is similar to grading of breast cancers according to histology [66]. This means that while there is a relative agreement of the tissue types that were operated in this study, a substantial degree of uncertainty remains, limiting the performance of the classifier as some samples are likely classified incorrectly by the annotator, giving inconsistent signals to the classification algorithm. This is a universal problem to methods that rely on machine learning, and we encourage authors to assess and report inter-observer agreements.

In addition to annotation, tissue heterogeneity and environmental factors are likely to play a role in the classification accuracy. We noted significant heterogeneity within the tissue classes and the issue has also been reported in other studies. For example, a recent REIMS study has shown that the molecular profile of stromal tissue is highly dependent on the distance from the tumor [67]. This means that by limiting the class division to four general classes, the classification problem suffers from high within-class variance, which ultimately affects the classification performance, when the DMS profiles of different classes share characteristics. The high within-class variance due to the tissue heterogeneity, further complicate the classification, when the number of available training samples is low. Due to the nature of the surgeries, some of the tissue classes are more common than others, which leads to disproportion between the classes and insufficiency of training samples to create a fully generalizable model. In addition, the regularized LDA classification might not be the optimal method for the identification of the tissues due to its simplicity, even though it has previously performed well in a more controlled setting. More complex and robust machine learning methods such as convolutional neural networks, could better compensate the heterogeneity of the samples and changes in environmental conditions, and work better if the number of training samples were increased to several thousands.

An additional source of within-class variance is the variation of environmental conditions between measurement days. The day-to-day variation has proven to be a limitation for the DMS-based system in previous studies, where the generalizability of the classification models has decreased, if the environmental conditions have varied between measurement sets [37], [38]. The operation room was assumed to be a more controlled environment in terms of humidity and temperature than a standard research lab. However, in this study, the variance in the operation room relative humidity was surprisingly high, ranging from 12% to 37% between operation days. The values do not match the recommendations of international standards. For example, the American Institute of Architects guidelines for Design and Construction of Hospital and Health Care Facilities recommend that the relative humidity should be between 30% to 60% [68].

To our knowledge, this was the first time a head-mounted system with gaze-tracking was used intraoperatively in operation room setting to help annotate operated tissues. The system provided a minimally intrusive method to record surgeries and although gaze-tracking was not used in the analysis of data, it could, in the future, be used to audit and study the work of different surgeons with varying levels of experience to see if their concentration on the operation area has variation and if there are general aspects in the operation room environment that can distract the operation.

In the future, emphasis should be placed on higher sampling frequency and annotation and enhanced control of environmental factors to reduce the day-to-day variation. With optimized hardware we expect significantly improved performance that could challenge that of MS. The improved system should subsequently be validated in a larger clinical trial, where a classification model for positive margins could also be created. If the performance matches or exceeds previous ex vivo results [39], the technology could be implemented to clinical practice to aid in margin assessment and to reduce avoidable reoperations.

V. CONCLUSION

In this study, we demonstrated the feasibility of intraoperative DMS-based tissue identification for the first time. We identified significant differences between the tissues operated during breast surgery. The use of the device did not prolong operation times or add complications. The results in tissue

identification do not yet warrant the use of the technology in a clinical application. There are multiple technical aspects of the system that can be improved, most significant of which is the measurement delays that can be overcome with a purpose-built DMS sensor that is better suited for real time measurements. In the future, the video-based annotation process should also be improved as the current inter-rater agreement was only moderate and likely had a significant decreasing effect on the classification accuracy.

ACKNOWLEDGMENT

The authors would like to thank Mikko Koskenranta for his assistance in software development for the sample preprocessing systems, Milena Pentti and Miia Helander for their assistance in the organizing of the surgeries, and all the helpful staff in Hatanpää Hospital. Anton Kontunen (part-time employee, shareholder), Markus Karjalainen (part-time employee, shareholder), Anna Anttalainen (employee at the time of writing), Pekka Kumpulainen (shareholder), Osmo Anttalainen (employee, shareholder), Niku Oksala (part-time employee, shareholder), and Antti Roine (employee, shareholder) are affiliated with 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. The study sponsors did not have any involvement in the study design; collection, analysis, and interpretation of data; the writing of the manuscript; or the decision to submit the manuscript for publication. *(Anton Kontunen and Ulla Karhunen-Enckell contributed equally to this work.)*

REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, 2018.

[2] Early Breast Cancer Trialists' Collaborative Group, "Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: An overview of the randomised trials," *Lancet*, vol. 366, no. 9503, pp. 2087–2106, 2005.

[3] Early Breast Cancer Trialists' Collaborative Group, "Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: Meta-analysis of individual patient data for 10 801 women in 17 randomised trials," *Lancet*, vol. 378, no. 9804, pp. 1707–1716, 2011.

[4] J. F. Waljee, E. S. Hu, P. A. Ubel, D. M. Smith, L. A. Newman, and A. K. Alderman, "Effect of esthetic outcome after breast-conserving surgery on psychosocial functioning and quality of life," *J. Clin. Oncol.*, vol. 26, no. 20, pp. 7–3331, Jul. 2008.

[5] J. Heil, E. Czink, M. Golatta, S. Schott, H. Hof, E. Jenetzky, M. Blumenstein, A. Maleika, G. Rauch, and C. Sohn, "Change of aesthetic and functional outcome over time and their relationship to quality of life after breast conserving therapy," *Eur. J. Surgical Oncol.*, vol. 37, no. 2, pp. 116–121, Feb. 2011.

[6] C. A. Garcia-Etienne, M. Tomatis, J. Heil, K. Friedrichs, R. Kreienberg, A. Denk, M. Kiechle, F. Lorenz-Salehi, R. Kimmig, G. Emons, M. Danaei, V. Heyl, U. Heindrichs, C. J. Rageth, W. Janni, L. Marotti, M. R. D. Turco, and A. Ponti, "Mastectomy trends for early-stage breast cancer: A report from the EUSOMA multi-institutional European database," *Eur. J. Cancer*, vol. 48, no. 13, pp. 1947–1956, Sep. 2012.

[7] M. Lautner, H. Lin, Y. Shen, C. Parker, H. Kuerer, S. Shaitelman, G. Babiera, and I. Bedrosian, "Disparities in the use of breast-conserving therapy among patients with early-stage breast cancer," *JAMA Surg.*, vol. 150, no. 8, pp. 778–786, Aug. 2015.

[8] M. Morrow, K. J. Van Zee, L. J. Solin, N. Houssami, M. Chavez-MacGregor, J. R. Harris, J. Horton, S. Hwang, P. L. Johnson, M. L. Marinovich, and S. J. Schnitt, "Society of surgical oncology—American society for radiation oncology—American society of clinical oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in ductal carcinoma *in situ*," *J. Clin. Oncol.*, vol. 34, no. 33, pp. 4040–4046, Aug. 2016.

[9] W. J. Gradishar, B. O. Anderson, J. Abraham, R. Aft, D. Agnese, K. H. Allison, S. L. Blair, H. J. Burstein, C. Dang, A. D. Elias, and S. H. Giordano, "Breast cancer, version 3.2020, NCCN clinical practice guidelines in oncology," *J. Nat. Comprehensive Cancer Netw.*, vol. 18, no. 4, pp. 452–478, 2020.

[10] F. Cardoso, S. Kyriakides, S. Ohno, F. Penault-Llorca, P. Poortmans, I. T. Rubio, S. Zackrisson, and E. Senkus, "Early breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up," *Ann. Oncol.*, vol. 30, no. 8, pp. 1194–1220, Aug. 2019.

[11] M. S. Moran, S. J. Schnitt, A. E. Giuliano, J. R. Harris, S. A. Khan, J. Horton, S. Klimberg, M. Chavez-MacGregor, G. Freedman, N. Houssami, P. L. Johnson, and M. Morrow, "Society of surgical oncology—American society for radiation oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in stages I and II invasive breast cancer," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 88, no. 3, pp. 553–564, Mar. 2014.

[12] L. Niinikoski, M. H. K. Leidenius, P. Vaara, A. Voynov, P. Heikkilä, J. Mattson, and T. J. Meretoja, "Resection margins and local recurrences in breast cancer: Comparison between conventional and oncoplastic breast conserving surgery," *Eur. J. Surgical Oncol.*, vol. 45, no. 6, pp. 976–982, Jun. 2019.

[13] J. Landercasper, A. J. Borgert, O. M. Fayanju, H. Cody, S. Feldman, C. Greenberg, J. Linebarger, B. Pockaj, and L. Wilke, "Factors associated with reoperation in breast-conserving surgery for cancer: A prospective study of American society of breast surgeon members," *Ann. Surgical Oncol.*, vol. 26, no. 10, pp. 3321–3336, Jul. 2019.

[14] D. R. G. Faverly, L. Burgers, P. Bult, and R. Holland, "Three dimensional imaging of mammary ductal carcinoma *in situ*: Clinical implications," *Semin. Diagn. Pathol.*, vol. 11, no. 3, pp. 193–198, Aug. 1994.

[15] T. J. Yeatman, A. B. Cantor, T. J. Smith, S. K. Smith, D. S. Reintgen, M. S. Miller, N. N. K. Ku, P. A. Baekey, and C. E. Cox, "Tumor biology of infiltrating lobular carcinoma: Implications for management," *Ann. Surgery*, vol. 222, no. 4, pp. 549–561, Oct. 1995.

[16] L. G. Melstrom, K. A. Melstrom, E. C. Wang, M. Pilewskie, and D. J. Winchester, "Ductal carcinoma *in situ*: Size and resection volume predict margin status," *Amer. J. Clin. Oncol.*, vol. 33, no. 5, pp. 438–442, Oct. 2010.

[17] M. M. Moore, G. Borossa, J. Z. Imbrie, R. E. Fechner, J. A. Harvey, C. L. Slingluff, R. B. Adams, and J. B. Hanks, "Association of infiltrating lobular carcinoma with positive surgical margins after breast-conservation therapy," *Ann. Surgery*, vol. 231, no. 6, pp. 877–882, Jun. 2000.

[18] N. Krekel, B. Zonderhuis, S. Müller, H. Bril, H.-J. van Slooten, E. de Lange de Klerk, P. van den Tol, and S. Meijer, "Excessive resections in breast-conserving surgery: A retrospective multicentre study," *Breast J.*, vol. 17, no. 6, pp. 602–609, Nov. 2011.

[19] N. B. Kouzminova, S. Aggarwal, A. Aggarwal, M. D. Allo, and A. Y. Lin, "Impact of initial surgical margins and residual cancer upon re-excision on outcome of patients with localized breast cancer," *Amer. J. Surgery*, vol. 198, no. 6, pp. 771–780, Dec. 2009.

[20] J. H. Volders, V. L. Negenborn, M. H. Haloua, N. M. A. Krekel, K. Jóźwiak, S. Meijer, and P. M. van den Tol, "Cosmetic outcome and quality of life are inextricably linked in breast-conserving therapy," *J. Surgical Oncol.*, vol. 115, no. 8, pp. 941–948, Jun. 2017.

[21] D. Q. Xue, C. Qian, L. Yang, and X. F. Wang, "Risk factors for surgical site infections after breast surgery: A systematic review and meta-analysis," *Eur. J. Surgical Oncol.*, vol. 38, no. 5, pp. 375–381, May 2012.

[22] Y. Grant, R. Al-Khudairi, E. St John, M. Barschkett, D. Cunningham, R. Al-Mufti, K. Hogben, P. Thiruchelvam, D. J. Hadjiminas, A. Darzi, A. W. Carter, and D. R. Leff, "Patient-level costs in margin re-excision for breast-conserving surgery," *Brit. J. Surgery*, vol. 106, no. 4, pp. 384–394, Feb. 2019.

[23] R. A. Cochrane, P. Valasiadou, A. R. M. Wilson, S. K. Al-Ghazal, and R. D. Macmillan, "Cosmesis and satisfaction after breast-conserving surgery correlates with the percentage of breast volume excised," *Brit. J. Surgery*, vol. 90, no. 12, pp. 1505–1509, Dec. 2003.

[24] S. L. Blair, K. Thompson, J. Rococco, V. Malcarne, P. D. Beitsch, and D. W. Ollila, "Attaining negative margins in breast-conservation operations: Is there a consensus among breast surgeons?" *J. Amer. College Surgeons*, vol. 209, no. 5, pp. 608–613, Nov. 2009.

[25] M. Thill, "MarginProbe: Intraoperative margin assessment during breast conserving surgery by using radiofrequency spectroscopy," *Expert Rev. Med. Devices*, vol. 10, no. 3, pp. 301–315, May 2013.

[26] F. Schnabel, S. K. Boolbol, M. Gittleman, T. Karni, L. Tafra, S. Feldman, A. Police, N. B. Friedman, S. Karlan, D. Holmes, and S. C. Willey, "A randomized prospective study of lumpectomy margin assessment with use of MarginProbe in patients with nonpalpable breast malignancies," *Ann. Surgical Oncol.*, vol. 21, no. 5, pp. 1589–1595, May 2014.

[27] T. M. Allweis, Z. Kaufman, S. Lelcuk, I. Pappo, T. Karni, S. Schneebaum, R. Spector, A. Schindel, D. Hershko, M. Zilberman, J. Sayfan, Y. Berlin, A. Hadary, O. Olsha, H. Paran, M. Gutman, and M. Carmon, "A prospective, randomized, controlled, multicenter study of a real-time, intraoperative probe for positive margin detection in breast-conserving surgery," *Amer. J. Surgery*, vol. 196, no. 4, pp. 483–489, Oct. 2008.

[28] S. J. Erickson-Bhatt, R. M. Nolan, N. D. Shemonski, S. G. Adie, J. Putney, D. Darga, D. T. McCormick, A. J. Cittadine, A. M. Zysk, M. Marjanovic, E. J. Chaney, G. L. Monroy, F. A. South, K. A. Cradock, Z. G. Liu, M. Sundaram, P. S. Ray, and S. A. Boppart, "Real-time imaging of the resection bed using a handheld probe to reduce incidence of microscopic positive margins in cancer surgery," *Cancer Res.*, vol. 75, no. 18, pp. 3706–3712, Sep. 2015.

[29] R. Li, P. Wang, L. Lan, F. P. Lloyd, C. J. Goergen, S. Chen, and J.-X. Cheng, "Assessing breast tumor margin by multispectral photoacoustic tomography," *Biomed. Opt. Exp.*, vol. 6, no. 4, pp. 1273–1281, 2015.

[30] M. Sans, J. Zhang, J. Q. Lin, C. L. Feider, N. Giese, M. T. Breen, K. Sebastian, J. Liu, A. K. Sood, and L. S. Eberlin, "Performance of the MasSpec pen for rapid diagnosis of ovarian cancer," *Clin. Chem.*, vol. 65, no. 5, pp. 674–683, May 2019.

[31] P. Saudemont, J. Quanico, Y. M. Robin, A. Baud, J. Balog, B. Fatou, D. Tierny, Q. Pascal, K. Minier, M. Pottier, and C. Focsa, "Real-time molecular diagnosis of tumors using water-assisted laser desorption/ionization mass spectrometry technology," *Cancer Cell*, vol. 34, no. 5, pp. 840–851, 2018.

[32] D. L. Phelps, J. Balog, L. F. Gildea, Z. Bodai, A. Savage, M. A. El-Bahrawy, A. V. Speller, F. Rosini, H. Kudo, J. S. McKenzie, R. Brown, Z. Takáts, and S. Ghaem-Maghami, "The surgical intelligent knife distinguishes normal, borderline and malignant gynaecological tissues using rapid evaporative ionisation mass spectrometry (REIMS)," *Brit. J. Cancer*, vol. 118, no. 10, pp. 1349–1358, May 2018.

[33] S. Mason, E. Manoli, L. Poynter, J. Alexander, P. Paizs, A. Adebesin, R. Goldin, A. Darzi, Z. Takats, and J. Kinross, "Mass spectrometry transanal minimally invasive surgery (MS-TAMIS) to promote organ preservation in rectal cancer," *Surgical Endoscopy*, vol. 34, no. 8, pp. 3618–3625, Aug. 2020.

[34] E. R. St John, J. Balog, J. S. McKenzie, M. Rossi, A. Covington, L. Muirhead, Z. Bodai, F. Rosini, A. V. M. Speller, S. Shousha, R. Ramakrishnan, A. Darzi, Z. Takats, and D. R. Leff, "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.*, vol. 19, no. 1, May 2017.

[35] G. A. Eiceman, Z. Karpas, and H. H. Hill, "Introduction to ion mobility spectrometry," in *Ion Mobility Spectrometry*, 3rd ed. Boca Raton, FL, USA: CRC Press, 2013, pp. 1–21.

[36] A. Kontunen, M. Karjalainen, J. Lekkala, A. Roine, and N. Oksala, "Tissue identification in a porcine model by differential ion mobility spectrometry analysis of surgical smoke," *Ann. Biomed. Eng.*, vol. 46, no. 8, pp. 1091–1100, Aug. 2018.

[37] A. Kontunen, J. Tuominen, M. Karjalainen, O. Anttalainen, T. Tolonen, P. Kumpulainen, M. Lepomäki, A. Vehkaoja, N. Oksala, and A. Roine, "Differential mobility spectrometry imaging for pathological applications," *Express Mol. Pathol.*, vol. 117, pp. 1–8, Dec. 2020.

[38] A. Kontunen, M. Karjalainen, A. Anttalainen, O. Anttalainen, M. Koskenranta, A. Vehkaoja, N. Oksala, and A. Roine, "Real time tissue identification from diathermy smoke by differential mobility spectrometry," *IEEE Sensors J.*, vol. 21, no. 1, pp. 717–724, Jan. 2021.

[39] M. Sutinen, A. Kontunen, M. Karjalainen, J. Kiiski, J. Hannus, T. Tolonen, A. Roine, and N. Oksala, "Identification of breast tumors from diathermy smoke by differential ion mobility spectrometry," *Eur. J. Surgical Oncol.*, vol. 45, no. 2, pp. 141–146, Feb. 2019.

[40] Finnish Breast Cancer Group. *RINTASYÖVÄN VALTAKUNNALLINEN DIAGNOSTIIKKA-JA HOITOSUOSITUS*. Accessed: Oct. 14, 2021. [Online]. Available: https://rintasyoparyhma.yhdistysavain.fi/hoitosuositus/

[41] W. J. Gradishar *et al.*, "Breast cancer, version 3.2020," *J. Nat. Compr. Cancer Netw.*, vol. 18, no. 4, pp. 452–478, 2020.

[42] M. Kassner, W. Patera, and A. Bulling, "Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput., Adjunct Publication*, Sep. 2014, pp. 1151–1160.

[43] R Core Team. *R: A Language and Environment for Statistical Computing*. Accessed: Oct. 30, 2021. [Online]. Available: https://www.r-project.org/

[44] RStudio Team. *RStudio: Integrated Development Environment for R*. Accessed: Oct. 30, 2021. [Online]. Available: https://www.rstudio.com/

[45] J. L. Fleiss and J. Cohen, "The equivalence of weighted Kappa and the intraclass correlation coefficient as measures of reliability," *Educ. Psychol. Meas.*, vol. 33, no. 3, pp. 613–619, 1973.

[46] M. A. Rotondi. (2018). *kappaSize: Sample Size Estimation Functions for Studies of Interobserver Agreement*. Accessed: Oct. 30, 2021. [Online]. Available: https://cran.r-project.org/package=kappaSize

[47] C. M. Bishop, "Linear models for classification," in *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006, pp. 179–196.

[48] T. A. Arnold and J. W. Emerson, "Nonparametric goodness-of-fit tests for discrete null distributions," *R J.*, vol. 3, no. 2, pp. 34–39, 2011.

[49] J. Alexander, L. Gildea, J. Balog, A. Speller, J. McKenzie, L. Muirhead, A. Scott, C. Kontovounisios, S. Rasheed, J. Teare, J. Hoare, K. Veselkov, R. Goldin, P. Tekkis, A. Darzi, J. Nicholson, J. Kinross, and Z. Takats, "A novel methodology for *in vivo* endoscopic phenotyping of colorectal cancer based on real-time analysis of the mucosal lipidome: A prospective observational study of the iKnife," *Surgical Endoscopy*, vol. 31, no. 3, pp. 1361–1370, Mar. 2017.

[50] R. J. Gray, B. A. Pockaj, E. Garvey, and S. Blair, "Intraoperative margin management in breast-conserving surgery: A systematic review of the literature," *Ann. Surgical Oncol.*, vol. 25, no. 1, pp. 18–27, Jan. 2018.

[51] N. M. Krekel, M. H. Haloua, A. M. Lopes Cardozo, R. H. de Wit, A. M. Bosch, L. M. de Widt-Levert, S. Müller, H. van der Veen, E. Bergers, E. S. de Lange de Klerk, S. Meijer, and M. P. van den Tol, "Intraoperative ultrasound guidance for palpable breast cancer excision (COBALT trial): A multicentre, randomised controlled trial," *Lancet Oncol.*, vol. 14, no. 1, pp. 48–54, Jan. 2013.

[52] M. Ahmed and M. Douek, "Intra-operative ultrasound versus wire-guided localization in the surgical management of non-palpable breast cancers: Systematic review and meta-analysis," *Breast Cancer Res. Treat.*, vol. 140, no. 3, pp. 435–446, 2013.

[53] H. Pan, N. Wu, H. Ding, Q. Ding, and J. Dai, "Intraoperative ultrasound guidance is associated with clear lumpectomy margins for breast cancer: A systematic review and meta-analysis," *PLoS ONE*, vol. 8, no. 9, p. 74028, 2013.

[54] V. S. Klimberg, "Advances in the diagnosis and excision of breast cancer," in *The American Surgeon*, vol. 69, no. 1. Atlanta, GA, USA: SAGE Publications, 2003, pp. 11–14.

[55] E. R. St John, R. Al-Khudairi, H. Ashrafian, T. Athanasiou, Z. Takats, D. J. Hadjiminas, A. Darzi, and D. R. Leff, "Diagnostic accuracy of intraoperative techniques for margin assessment in breast cancer surgery: A meta-analysis," *Ann. Surgery*, vol. 265, no. 2, pp. 300–310, 2017.

[56] M. Thill, C. Dittmer, K. Baumann, K. Friedrichs, and J.-U. Blohmer, "MarginProbe—Final results of the German post-market study in breast conserving surgery of ductal carcinoma *in situ*," *Breast*, vol. 23, no. 1, pp. 94–96, Feb. 2014.

[57] K. Ojala, T. J. Meretoja, J. Mattson, P. Salminen-Peltola, S. Leutola, M. Berggren, and M. H. K. Leidenius, "The quality of preoperative diagnostics and surgery and their impact on delays in breast cancer treatment—A population based study," *Breast*, vol. 26, pp. 80–86, Apr. 2016.

[58] M. E. Taylor, C. A. Perez, K. J. Halverson, R. R. Kuske, G. W. Philpott, D. M. Garcia, J. E. Mortimer, R. J. Myerson, D. Radford, and C. Rush, "Factors influencing cosmetic results after conservation therapy for breast cancer," *Int. J. Radiat. Oncol. Biol. Physics*, vol. 31, no. 4, pp. 753–764, Feb. 1995.

[59] E. Parvez, S. D. Cornacchi, N. Hodgson, A. Thoma, I. Kong, G. Foster, J. Cheng, C. H. Goldsmith, D. Dao, and P. J. Lovrics, "A cosmesis outcome substudy in a prospective, randomized trial comparing radioguided seed localization with standard wire localization for nonpalpable, invasive, and *in situ* breast carcinomas," *Amer. J. Surgery*, vol. 208, no. 5, pp. 711–718, Nov. 2014.

[60] T. Hashem, A. Morsi, A. Farahat, T. Zaghloul, and A. Hamed, "Correlation of specimen/breast volume ratio to cosmetic outcome after breast conserving surgery," *Indian J. Surgical Oncol.*, vol. 10, no. 4, pp. 668–672, Dec. 2019.
[61] P. Panhofer, V. Ferenc, M. Schütz, A. Gleiss, P. Dubsky, R. Jakesz, M. Gnant, and F. Fitzal, "Standardization of morbidity assessment in breast cancer surgery using the clavien dindo classification," *Int. J. Surgery*, vol. 12, no. 4, pp. 334–339, Apr. 2014.
[62] J. Edwards, D. Mcmillan, S. Stallard, J. Doughty, L. Romics, and F. Savioli, "The effect of postoperative complications on survival and recurrence after surgery for breast cancer: A systematic review and meta-analysis," *Eur. J. Cancer*, vol. 138, p. S108, Oct. 2020.
[63] A. Lucci, L. M. McCall, P. D. Beitsch, P. W. Whitworth, D. S. Reintgen, P. W. Blumencranz, A. M. Leitch, S. Saha, K. K. Hunt, and A. E. Giuliano, "Surgical complications associated with sentinel lymph node dissection (SLND) plus axillary lymph node dissection compared with SLND alone in the American college of surgeons oncology group trial Z0011," *J. Clin. Oncol.*, vol. 25, no. 24, pp. 3657–3663, Aug. 2007.
[64] M. Donker *et al.*, "Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer (EORTC 10981–22023 AMAROS): A randomised, multicentre, open-label, phase 3 non-inferiority trial," *Lancet Oncol.*, vol. 15, no. 12, pp. 1303–1310, Nov. 2014.
[65] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, Mar. 1977.
[66] P Boiesen, P. O. Bendahl, L. Anagnostaki, H. Domanski, E. Holm, I. Idvall, S. Johansson, O. Ljungberg, and A. Ringberg, "Histologic grading in breast cancer: Reproducibility between seven pathologic departments," *Acta Oncol.*, vol. 39, no. 1, pp. 41–45, Jan. 2000.
[67] P.-M. Vaysse, L. F. S. Kooreman, S. M. E. Engelen, B. Kremer, S. W. M. O. Damink, R. M. A. Heeren, M. L. Smidt, and T. P. Siegel, "Stromal vapors for real-time molecular guidance of breast-conserving surgery," *Sci. Rep.*, vol. 10, no. 1, pp. 1–6, Dec. 2020.
[68] *AIA Guidelines for Design and Construction of Hospital and Health Care Facilities*, Facility Guidelines Inst., Amer. Inst. Architects, Washington, DC, USA, 2006.

ANTON KONTUNEN received the B.Sc. and M.Sc. degrees in bioengineering from the Tampere University of Technology, Finland, in 2015 and 2017, respectively. He is currently pursuing the D.Sc. degree with the Faculty of Medicine and Health Technology, Tampere University, Finland.

From 2014 to 2017, he was a Research Assistant with the Department of Automation Science and Engineering, Tampere University of Technology. He works as a System Engineer in a university spin-off company, Olfactomics Oy, Tampere, Finland. His research interests include biomedical engineering, sensor technology, and data science. He was a recipient of the Professional Community of Academic Engineers and Architects in Finland (TEK) Award for the best master's thesis in Finland, in 2017.

ULLA KARHUNEN-ENCKELL received the M.D. degree in medicine from Southwestern University, Finland, in 1998.

Since 2008, she has been a Consultant Plastic Surgeon. She is currently the Chief of the Breast Surgery Unit, Tampere University Hospital. Her research interests include surgical oncology, breast surgery from plastic surgeon's perspective, and sensor technology. She serves as a Board Member for the Finnish Breast Surgery Association.

MARKUS KARJALAINEN received the M.Sc. degree in automation engineering from the Tampere University of Technology, Finland, in 2007. He is currently pursuing the D.Sc. degree with the Faculty of Medicine and Health Technology, Tampere University, Finland.

From 2006 to 2017, he was with the Department of Automation Science and Engineering, Tampere University of Technology. He works as a Hardware Engineer in a university spin-off company, Olfactomics Oy, Tampere, Finland. His research interests include chemical communication and sensor technology.

ANNA ANTTALAINEN received the B.Sc. degree in bioinformation technology (major in computational and cognitive biosciences) and the M.Sc. degree in life science technologies (major in complex systems) from Aalto University, Finland, in 2016 and 2019, respectively.

From 2014 to 2016, she was a Research Assistant with the Department of Neuroscience and Biomedical Engineering. From 2016 to 2019, she was an Assistant Teacher for a mathematical modeling course with Aalto University. From 2019 to 2021, she worked as a Data Scientist at Olfactomics Oy. She is currently pursuing her career in data science at Medaffcon Oy, Espoo, Finland. Her research interests include machine learning with a focus on neural networks.

PEKKA KUMPULAINEN received the M.Sc. and D.Sc. (Tech.) degrees from the Tampere University of Technology, in 1994 and 2014, respectively.

From 1992 to 2019, he was with the Department of Automation Science and Engineering, Tampere University of Technology. Since 2019, he has been working as a Data Analysis Engineer with Olfactomics Oy and an independent consultant in data analytics.

LEENA PITKÄNEN received the M.D. degree in medicine from the University of Eastern Finland, in 2011.

She has been working as a Consultant Plastic Surgeon at the Breast Unit, Tampere University Hospital, since 2018. She has also been the Chairperson of the Local Chapter of Junior Doctors' Association, Finland, and the Young Plastic Surgeons' Association, Finland. Her special scientific research interests include lipotransplantation of the breast, breast conserving surgery, and breast reconstructions.

OSMO ANTTALAINEN received the M.Sc. degree in energy technology from the Lappeenranta University of Technology, Finland, in 1992.

From 1994 to 2018, he was employed by Environics Oy, Mikkeli, Finland. Currently, he is continuing his career in a university spin-off company, Olfactomics Oy, Tampere, Finland. His research interests include ion mobility spectrometry and mixed signal electronics.

ANTTI VEHKAOJA (Member, IEEE) received the D.Sc. (Tech.) degree in automation science and engineering from the Tampere University of Technology, Tampere, Finland, in 2015.

He is currently an Associate Professor (tenure track) of sensor technology and biomeasurements at the Faculty of Medicine and Health Technology, Tampere University, Finland. He has authored more than 100 scientific articles, mainly in the area of biomedical engineering. His research interests include embedded measurement technologies for physiological monitoring and related signal processing and data analysis methods as well as analysis of volatile organic compounds in biomedical applications.

NIKU OKSALA received the M.D. and Ph.D. degrees in medicine and experimental surgery from the University of Eastern Finland, in 2000 and 2003, respectively, and the D.Sc. (Med.) degree in molecular biology from Tampere University, Finland, in 2009.

Since 2007, he has been a Consultant Vascular Surgeon and a Clinical Teacher, and since 2014, has also been an Associate Professor of surgery being a tenured Full Professor, in early 2018. He is currently a Professor of vascular surgery with the Faculty of Medicine and Health Technology, Tampere University, and the Chief Vascular Surgeon with the Tampere University Hospital. He is the chairperson of the board and the CMO of Olfactomics Oy, Tampere, Finland. He has authored more than 190 international journal articles. His current research interests include biomedical sensor technology, clinical research, and molecular biology. He serves as a Board Member for the Instrumentarium Science Foundation and the Finnish Cardiovascular Research Center, Tampere, Finland.

ANTTI ROINE received the M.D. and Ph.D. degrees from Tampere University, Finland, both in 2014.

Since 2014, he has held various positions as a Physician in the field of surgery, emergency medicine, and family medicine. In this period, he has also acted as a Project Manager in healthcare digitalization in the city of Tampere. Currently, he acts as the CEO of Olfactomics Oy and is an active Researcher at Tampere University. His research interests include surgical oncology, gas sensors, and application of artificial intelligence in medical applications.