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.

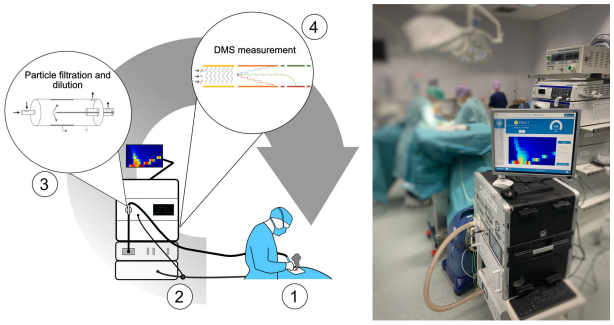

FIGURE 1. Simplified operation principle of the measurement system (left) and the measurement system in an operation room (right). 1) Surgical smoke is produced by the diathermy instrument. 2) Induced current from the dispersive electrode triggers the measurement. 3) Surgical smoke sample is taken into 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.

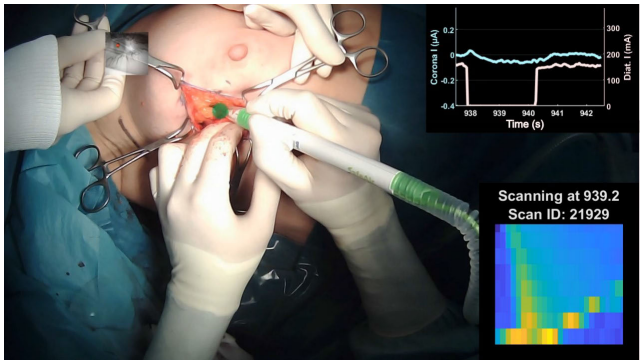

FIGURE 2. Still image of a video file that was used in tissue annotation. The DMS spectrum of the measurement is overlayed in the bottom right corner and the current of the electric corona filter and the measured induction current from the dispersive diathermy electrode can be seen in the upper right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

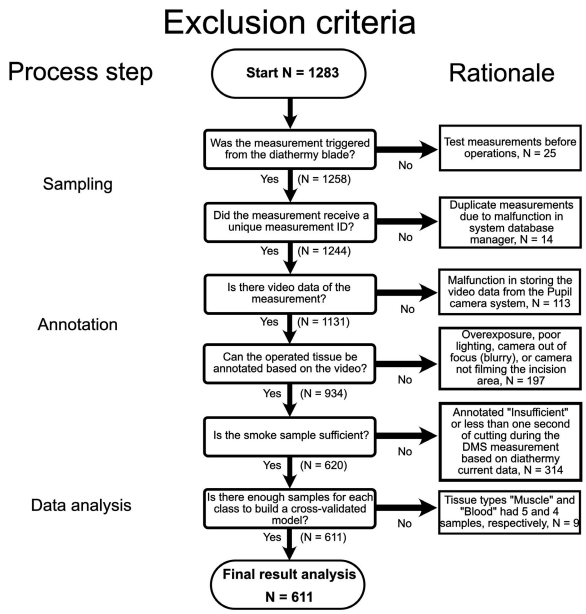

FIGURE 3. The exclusion criteria for the DMS measurement data.

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.

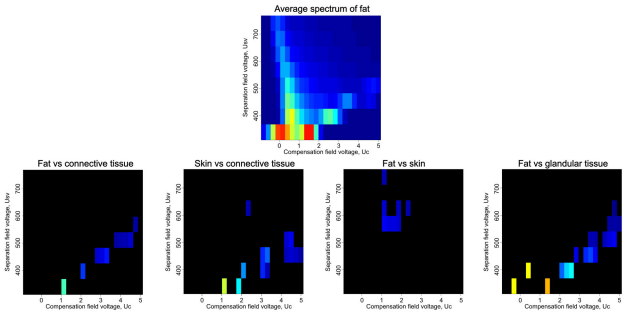

FIGURE 4. An average DMS spectrum of fat alongside features of the spectra that have a statistically significant difference between the tissue classes.

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.