

<https://helda.helsinki.fi>

Detection of cultured breast cancer cells from human tumor-derived matrix by differential ion mobility spectrometry

Lindfors, Lydia

2022-04-15

Lindfors , L , Sioris , P , Anttalainen , A , Korelin , K , Kontunen , A , Karjalainen , M , Naakka , E , Salo , T , Vehkaoja , A , Oksala , N , Hytönen , V , Roine , A & Lepomäki , M 2022 , ' Detection of cultured breast cancer cells from human tumor-derived matrix by differential ion mobility spectrometry ' , Analytica Chimica Acta , vol. 1202 , 339659 . <https://doi.org/10.1016/j.aca.2022.339659>

<http://hdl.handle.net/10138/342956>

<https://doi.org/10.1016/j.aca.2022.339659>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.



Detection of cultured breast cancer cells from human tumor-derived matrix by differential ion mobility spectrometry



Lydia Lindfors^{a,1}, Patrik Sioris^{a,1,**}, Anna Anttalainen^b, Katja Korelin^{c,d}, Anton Kontunen^{a,b}, Markus Karjalainen^{a,b}, Erika Naakka^{c,d}, Tuula Salo^{c,d}, Antti Vehkaoja^a, Niku Oksala^{a,b,e}, Vesa Hytönen^{a,f}, Antti Roine^b, Maiju Lepomäki^{a,f}

^a Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

^b Olfactomics Ltd., Tampere, Finland

^c Department of Oral and Maxillofacial Diseases, University of Helsinki, 00014, Helsinki, Finland

^d Translational Immunology Research Program (TRIMM), University of Helsinki, 00014, Helsinki, Finland

^e Vascular Centre, Tampere University Hospital, Tampere, Finland

^f Fimlab Laboratories, Tampere, Finland

HIGHLIGHTS

- Differential ion mobility spectrometry is an applicable method for tissue analysis.
- Human myoma-based matrix can be used to mimic tumor microenvironment in *in vitro* studies.
- Low densities of cultured breast cancer cells can be detected from Myogel with DMS.

ARTICLE INFO

Article history:

Received 21 December 2021

Received in revised form

22 February 2022

Accepted 25 February 2022

Available online 28 February 2022

ABSTRACT

The primary treatment of breast cancer is the surgical removal of the tumor with an adequate healthy tissue margin. An intraoperative method for assessing surgical margins could optimize tumor resection. Differential ion mobility spectrometry (DMS) is applicable for tissue analysis and allows for the differentiation of malignant and benign tissues. However, the number of cancer cells necessary for detection remains unknown. We studied the detection threshold of DMS for cancer cell identification with a widely characterized breast cancer cell line (BT-474) dispersed in a human myoma-based tumor microenvironment mimicking matrix (Myogel). Predetermined, small numbers of cultured BT-474 cells were dispersed into Myogel. Pure Myogel was used as a zero sample. All samples were assessed with a DMS-based custom-built device described as “the automated tissue laser analysis system” (ATLAS). We used machine learning to determine the detection threshold for cancer cell densities by training binary classifiers to distinguish the reference level (zero sample) from single predetermined cancer cell density levels. Each classifier (sLDA, linear SVM, radial SVM, and CNN) was able to detect cell density of 3700 cells μL^{-1} and above. These results suggest that DMS combined with laser desorption can detect low densities of breast cancer cells, at levels clinically relevant for margin detection, from Myogel samples *in vitro*.

© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Breast cancer is the leading cause of women's cancer-related deaths [1]. The cornerstone of breast cancer treatment is tumor

removal preferably with breast-conserving surgery (BCS), to which approximately 80% of patients are amenable [2,3]. Achieving healthy tissue margins is essential, as margin involvement increases the risk of local recurrence [4]. In invasive carcinoma, the

** Corresponding author.

E-mail address: patrik.sioris@tuni.fi (P. Sioris).

¹ Authors contributed equally.

tissue margin is adequate if the ink-marked tissue edge of the surgical specimen is free from cancer. A 2 mm margin is recommended for ductal carcinoma *in situ* (DCIS) [2]. In case of margin positivity, re-excision is usually required [2,4,5]. Approximately 30% of patients treated with BCS require re-excision [6]. However, re-excisions impair the aesthetic results of BCS [7] and increase the burden on the health care system.

A tool for margin assessment during surgery could optimize the resection of the tumor and reduce the incidence of positive margins. Several methods for intraoperative tissue identification have been developed, but they suffer from resource intensiveness and lack of specificity or sensitivity [8]. DCIS poses a specific challenge as this precursor of invasive cancer grows along the mammary ducts, and the cancer cell density at the resection margin can be low. Additionally, the assessment of DCIS with frozen section analysis can be challenging due to its diffuse growth pattern and can result in undetected positive margins on the conventional breast pathology pipeline [9].

Molecular methods for tissue analysis have shown promise. Most emerging technologies utilize mass spectrometry (MS) [10–12]. Although MS is unparalleled in accuracy, its use is limited by its cost and complexity. Differential Mobility Spectrometry (DMS) is an atmospheric technology in which the gaseous molecules of the sample are filtered, ionized, and derived into a high voltage asymmetric electric field. The trajectories of the molecules differ according to their shape, size, and charge. The separated ions hit a detector plate and lose their charge which induces an electric current measured by the device. The molecular composition of the sample is then presented as a dispersion plot. The advantages of DMS are the low complexity of analysis, easy hardware maintenance, and economic efficiency compared to MS. Our research team has developed a DMS-based method for tissue analysis, in which the sampling is conducted by electrocautery or laser evaporation. The method has demonstrated promising results in studies on animal tissues, human breast cancer, and brain tumors [13–17].

Although we have demonstrated the ability to differentiate malignant and benign tissues with DMS, the number of cancer cells necessary for detection remains undecided. The study of the matter poses a challenge as the volume of cancer cells in surgical specimens varies between 15% and 95% [18]. In addition, the cell contents of surgical specimens can be little, and the growth patterns of cells diffuse. The heterogeneous composition of the matrix is also a significant confounder in *ex vivo* specimens. An approach that utilizes cultured cells and a tumor microenvironment (TME) mimicking matrix is appealing as the matrix remains similar throughout analysis and the number of cells in the matrix can be accurately controlled.

The aim of this paper is to determine the detection threshold for breast cancer cells with laser desorption DMS with a widely characterized breast cancer cell line and a human leiomyoma-based TME matrix (Myogel). In future, this study design could act as a reproducible and standardizable means for comparing tissue analysis methods and their calibration.

2. Materials & methods

2.1. Chemicals and reagents

We purchased the Dulbecco's Modified Eagle's Medium (DMEM), antibiotic mixture (10 000 U penicillin and 10 mg streptomycin per ml in 0.9% NaCl), 10% Dimethyl Sulfoxide (DMSO), Dulbecco's Phosphate-buffered saline (DPBS), and aprotinin (66.7 mg mL⁻¹) from Sigma-Aldrich (St. Louis, MO, USA). Fetal bovine serum (FBS) was purchased from Gibco (Waltham, Massachusetts, USA). Thrombin (0.6 U mL⁻¹) and fibrinogen (1 mg mL⁻¹)

were purchased from Merck (Taufkirchen, Germany). Myogel (1 mg mL⁻¹) was acquired from the University of Oulu (Oulu, Finland).

2.2. Cell line and Myogel matrix preparation

We utilized the breast cancer cell line BT-474 of which the intrinsic subtype is Luminal B-like (HER2-positive) [19]. The BT-474 cells were cultured in DMEM with low glucose and supplemented with 10% FBS and an antibiotic mixture. The cells were incubated in a Thermo Scientific CO₂ incubator (37 °C; 5% CO₂) and passaged every 2–3 days. Last, the cultured BT-474 cells were frozen in 10% DMSO.

Myogel matrix mixture was prepared as previously described by Salo et al. [20]. The BT-474 cells were defrosted, washed with DPBS, and centrifuged. We counted the cells (Countess™ automated cell counter, Invitrogen, cat no C10281) and suspended them with a culture media to densities ranging between 10² and 10⁶ cells in each 27 μL sample. The cell densities approximately corresponded to 3.7–37 000 cells μL⁻¹. We added the Myogel, thrombin, and aprotinin into these solutions simultaneously, and fibrinogen right before pipetting as it quickly forms a gel. Finally, each sample (27 μL) consisted of BT-474 breast cancer cells (10²–10⁶ cells per sample); DMEM (19.5 μL); Myogel (2.65 μL); and the above-described reagents (4.86 μL). We also prepared a cancer cell-free Myogel matrix mixture which performed as a zero sample.

The use of human leiomyoma in the preparation of Myogel has been approved by the Ethics Committees of Oulu University Hospital (statement number 2/2017). All methods were performed in accordance with the relevant guidelines and regulations.

For this study, we prepared eight custom-made well plates that each contained 36 wells (3.5 mm × 3.5 mm × 3.0 mm). Of each prepared cell density, 4–8 samples were pipetted into the wells in random order (Sartorius Mline® electronic pipette, Sartorius AG, Germany). In total, we prepared 87 samples. The remaining wells were filled with agar which produces an unnoticeable DMS response. These agar samples functioned as a quality control sample. After sample preparation, the well plates were placed in a cell culture incubator (37 °C) for 30 min for the Myogel matrix to solidify. A parafilm was placed on top of each well plate to prevent sample dehydration, and the samples were frozen (–20 °C).

2.3. Measurement system

The study utilized a custom-built automated tissue laser analysis system (ATLAS), which has been previously described in detail by Vehkaoja et al. [21] Contrary to the previous publication, the DMS sensor utilized in this study was a prototype version of the IonVision DMS device (Olfactomics Oy, Finland).

ATLAS comprised a height-adjustable sampling stage, laser evaporator, and the DMS device. Sample evaporation was performed with a computer controlled, 40 W, 10.6 μm laser cutter (CO₂ Laser engraving machine, Vevor, China). We replaced the control electronics and software of the laser with custom-built hardware and software to increase the spatial accuracy and reliability of sampling. Laser sampling was controlled by a graphical user interface. To control sample humidity and ensure the absence of volatile contaminants from room air, carrier gas of purified and humidified pressurized air was continuously directed to the sampling stage through a sampling nozzle. The nozzle created a protective stream of carrier gas around the sampling area, and upon sample vaporization, guided the sample gas to the DMS inlet tube. A schematic representation of the measurement system is introduced in Fig. 1 (see Fig. 2).

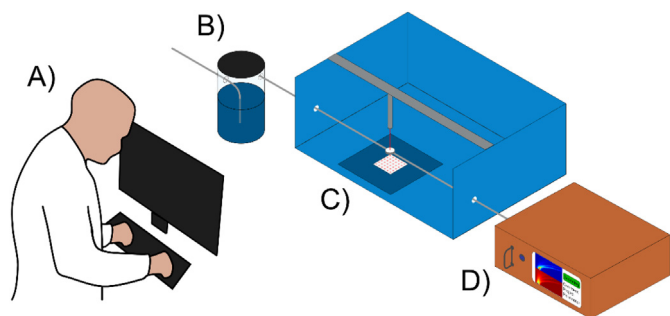


Fig. 1. The ATLAS measurement system: A) the graphical user interface; B) the air humidifier for the carrier gas; C) the sampling platform; and D) the DMS sensor.

2.4. Sampling

We analyzed each well plate within a period of two months on separate days to avoid day-wise bias or drift which could affect the results. Each sample (27 μL) was evaporated four times with the laser. The laser incisions were in a square shape 1 mm apart from each other vertically and horizontally. A 2 ns laser pulse was repeated 120 times with a pulse repetition time set to 98 ns. The start delay for the measurement was 1 s to account for the transfer of the sample from the sampling platform to the DMS core. We

implemented a cleansing period of 30 s after each measurement to reduce the carry-over from consecutive measurements. The vaporized samples were analyzed at compensation voltage values from -0.5 to 7 V with increments of 0.107 V, and separation voltage values of 200 – 800 V with increments of 15 V. Thus, the resulting dispersion plots had 2800 features for both the positive and negative sides, representing the molecular composition of the sample. With the DMS gap width of the IonVision instrument (0.25 mm), the voltage values for the compensation field and the separation field were equal to electric field strengths of -2 – 28 kV m^{-1} and 800 – 3200 kV m^{-1} , respectively.

2.5. Approximation of cell volume

We approximated the number of breast cancer cells that reached the DMS analysis computationally. The formula for the calculation of the volume percentages of breast cancer cells per well is presented in Table 1. The mean diameter of a viable BT-474 cell was 14.6 μm . Thus, the mean volume of a BT-474 cell was 1.63×10^{-6} μL . We calculated the volume percentage of BT-474 cells per well as follows: $10^n \times 1.63 \times 10^{-6}$ $\mu\text{L} \times (27 \mu\text{L})^{-1} \times 100\%$. The volume of a laser beam was 0.432 μL (diameter 0.25 mm; height 2.2 mm). Each well was measured four times. The estimated number of cells that reached the DMS analysis was calculated as follows: $10^n \times 4 \times 0.432 \mu\text{L} \times (27 \mu\text{L})^{-1}$.

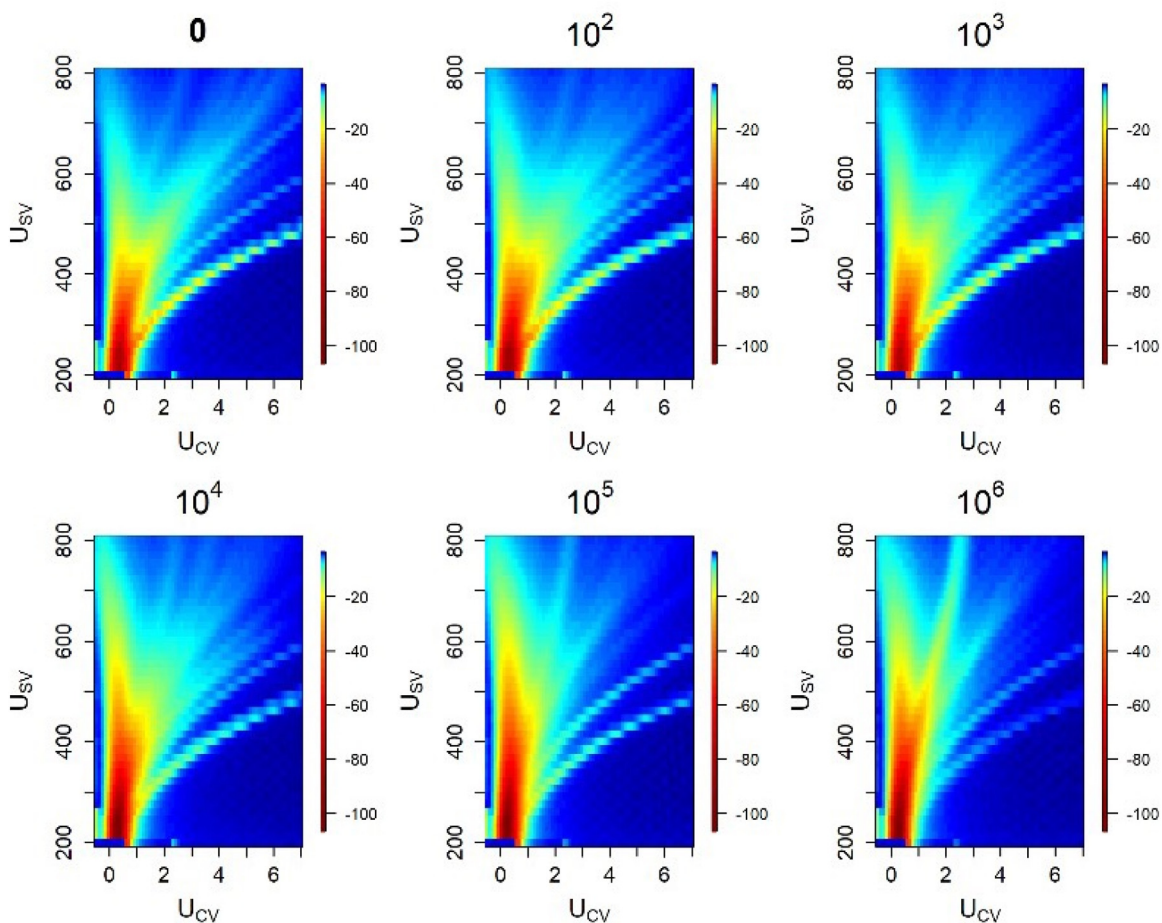


Fig. 2. The averaged DMS dispersion plots of the zero sample and the applied cell density levels. We plotted only the measurements of the positive ion mode. On high density levels, especially 10^5 and 10^6 , there is a visible, pronounced peak at the high U_{SV} values between 2 V and 3 V U_{CV} .

Table 1
The description of samples.

Number of breast cancer cells per well	Equivalent cell density (cells μL^{-1})	Volume percentage of breast cancer cells per well (%)	Estimated number of cells that reached DMS analysis
0	0	0	0
10^2	3.7	0.00064	6
10^3	37	0.0064	64
10^4	370	0.064	640
10^5	3700	0.64	6400
10^6	37 000	6.4	64 000

2.6. Statistical analysis

Raw DMS data accumulated from the positive ion mode were used in the statistical analysis. As the measurements of each well plate formed a time series, we performed dimension-wise linear trend removal for each measurement series prior to the statistical analysis to compensate for sensor drifting.

The statistical analysis was performed with the statistical software R [22] in the RStudio environment [23]. In the modelling, we utilized the R packages pls [24]; sda [25]; e1071 [26]; and keras [27] with Tensorflow backend [28].

2.6.1. Classification models

We created classification models for each density level and zero sample. Four different classifiers were employed: shrinkage linear discriminant analysis (sLDA); support vector machine (SVM) with linear kernel; SVM with radial kernel; and convolutional neural networks (CNNs). The lowest cell density that was reliably identified from the zero sample determined the estimated level of detection.

In LDA, the linear separability between classes is maximized by creating new linear combinations of the original data dimensions and employing a subset of them in the classification model. For high dimensional data sets, shrinkage regularization is often required [29]. sLDA has been previously applied in the classification of DMS data [16,17].

SVMs seek hyperplanes that best separate the classes. The data points which are the nearest to the hyperplanes are called support vectors. The hyperplanes are chosen to maximize the gap between classes defined by the support vectors. If the data is not linearly separable, it can be transformed to a higher-dimensional space with the so-called kernel trick, which enables non-linear classification [30]. In this study, we utilized both linear SVM and non-linear SVM with radial kernel.

CNNs are complex non-linear models specialized in the analysis of image-like data such as dispersion matrices. Neural networks consist of interconnected layers of computational units, which produce the target value as a hierarchical function of the inputs. The network parameters are updated iteratively to match the model's output to the true classes. Unlike other methods, CNNs can utilize spatial feature extraction from dispersion matrices [31]. CNNs have proven useful in DMS data analysis provided that the amount of training data is sufficient. We have piloted the use of CNNs in DMS data classification [15,32] and regression [33]. In this study, the CNN comprised two convolutional layers with fifteen and thirty 5×5 and 3×3 kernels, respectively, and 2×2 max pooling layers followed by two hidden layers with respective fifty and ten nodes. Tanh was used as the activation function. L1 regularization, batch normalization, and dropout with a 0.5 dropout rate were used to regularize the model.

2.6.2. Cross validation

Cross validation (CV) estimates the generalizability performance of a statistical model. In CV, the data is divided into k sets, each of

which is left out as an unseen test set at a time. The training data is used to construct a statistical model, whereupon the test data set is classified with. The results are reported as the combined results of all test folds.

Typically, CV is performed on randomly divided data sets. However, if the measurements are not independent, random split will lead to biased and overly optimistic results. In this study, each well plate was measured on a separate day and each well was measured four times repeatedly. The dynamic features related to measurement day can fluctuate. Therefore, the sample spectra from each well plate are more likely to resemble each other compared to those analyzed on a separate occasion. An additional confounding factor to sample independence is the carry-over of smoke from subsequent measurements. To reduce the risk of overfitting, we applied both well-wise and day-wise CV. The cross-validation methods are illustrated in Fig. 3.

2.6.3. Regression

Regression models a continuous output variable (e.g., density) as a function of input variables (e.g., the intensity of a DMS dispersion plot). In this study, we used partial least squares (PLS) regression to model the cellular densities. This is a linear regression technique in which the data dimensionality is reduced similarly to in principal component analysis. The application of PLS regression to DMS data has been described in detail [33]. In this study, the PLS regression model to predict the cellular density of each sample was cross-validated over measurement days and wells.

3. Results

We collected data on 348 Myogel-cell-sample measurements. Of these measurements, 86 were excluded due to air bubble formation during sample preparation; dehydration of the sample; and technical errors in the measurement device such as faulty dispersion plots. Thus, 262 measurements were included in the analysis. Agar-filled wells were used for calibration and were also excluded. Due to exclusions, the number of samples from each density level were not equal. The distribution of measurements according to measurement day and density level is shown in Table 2.

The threshold for the detection of cancer cells was estimated by training binary classifiers to distinguish the zero sample from each cell density. A particular cell density was considered detectable if the CV classification accuracy exceeded the theoretical no-information rate with statistical significance. Because the classes were uneven, the guess level was defined as the correct classification rate achieved by assigning all samples to the most prevalent class.

The classification accuracies and their corresponding McNemar's test p-values are presented in Table 3 and Table 4 and visualized in Fig. 4. As we utilized four models, to achieve a 0.05 false discovery rate, the cut-off limits for the p-values were Bonferroni corrected to $0.05/4 = 0.0125$. Omitting sLDA, we were able to detect cell densities of 10^5 cells/27 μL and higher with each classification model in both CV approaches. The Linear SVM and CNN models

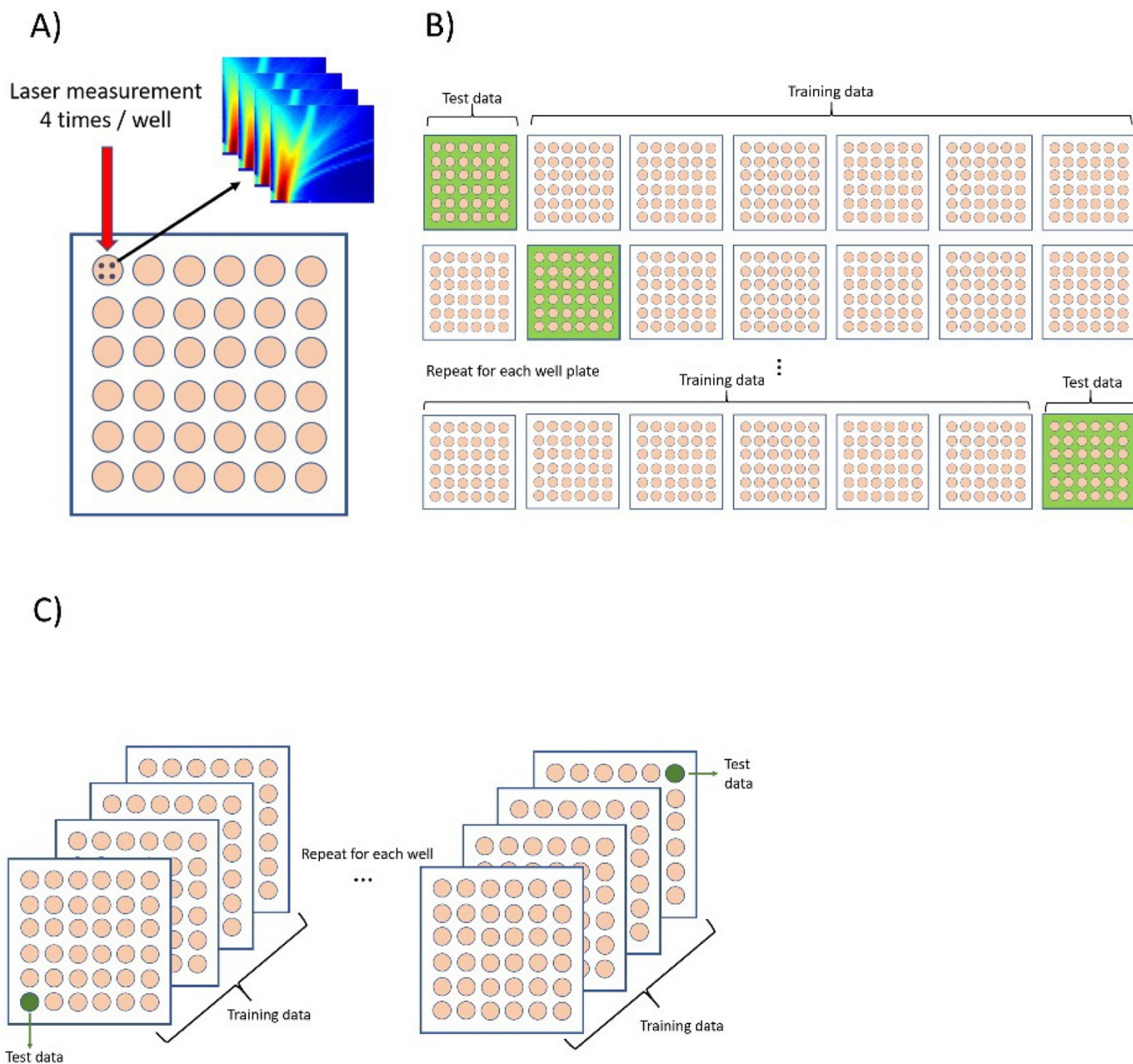


Fig. 3. A) We measured each well four times. B) Day-wise cross-validation (CV): each well plate was measured on a separate day and the data gathered during each day performed as the test set at a time. C) Well-wise CV: each well was left out as the test set at a time.

Table 2

The distribution of samples according to cell density level and measurement day.

Cell density	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Number (n)
0	0	4	0	0	4	11	12	0	31
10^2	4	0	4	4	8	11	8	0	39
10^3	0	4	4	4	2	4	8	8	34
10^4	16	0	0	16	4	10	11	3	60
10^5	0	20	4	0	7	7	8	8	54
10^6	0	4	12	0	8	8	4	8	44
Number (n)	20	32	24	24	33	51	51	27	262

achieved the respective detection levels of 10^4 and 10^3 cells in well-wise CV.

Next, we applied PLS regression to predict cancer cell densities. We used a base 10 logarithm of each cell density. The

predictions were scaled back to the original scale after analysis. The PLS regression results are shown in Fig. 5. In general, the measurements with lower cell densities were predicted close to the guess level. Even though we observed a slight elevation in

Table 3

The accuracies (Acc) obtained with day-wise cross validation (CV). A statistically significant p-value (P) was determined as the Bonferroni corrected significance level of 0.0125.

Cell density	10 ²		10 ³		10 ⁴		10 ⁵		10 ⁶	
Model	Acc	P	Acc	P	Acc	P	Acc	P	Acc	P
sLDA	51%	0.80	62%	0.085	66%	0.55	68%	0.22	85%	5.8×10^{-7}
Linear SVM	51%	0.80	55%	0.36	77%	0.016	80%	7.7×10^{-4}	91%	8.5×10^{-10}
Radial SVM	60%	0.27	65%	0.031	74%	0.073	78%	3.8×10^{-3}	89%	5.2×10^{-9}
CNN	53%	0.73	69%	4.2×10^{-3}	69%	0.29	81%	3.2×10^{-4}	95%	1.4×10^{-12}

Table 4

The accuracies (Acc) obtained with well-wise cross validation (CV). A statistically significant p-value (P) was determined as the Bonferroni corrected significance level of 0.0125.

Cell density	10 ²		10 ³		10 ⁴		10 ⁵		10 ⁶	
Model	Acc	P	Acc	P	Acc	P	Acc	P	Acc	P
sLDA	57%	0.45	54%	0.45	71%	0.16	81%	3.1×10^{-4}	89%	5.2×10^{-9}
Linear SVM	54%	0.64	60%	0.13	81%	9.1×10^{-4}	92%	2.1×10^{-9}	96%	1.1×10^{-13}
Radial SVM	61%	0.20	58%	0.19	75%	0.046	87%	1.1×10^{-6}	93%	1.4×10^{-11}
CNN	51%	0.80	75%	1.1×10^{-4}	81%	9.1×10^{-4}	88%	2.8×10^{-7}	95%	1.4×10^{-12}

the 10⁵ and 10⁶ predictions, these predictions are drawn towards guess level.

4. Discussion

In this study, we have demonstrated the capability of a laser-based DMS system to detect breast cancer cells from a Myogel matrix at density levels that are relevant to surgical margin detection. We detected cell densities of 3700 cells μL^{-1} and higher with DMS. The absolute number of evaporated cells required for detection was approximately 6400 cells/measurement. We achieved the highest classification accuracies with SVM and CNN. We also represented a PLS regression model which correlated to the cell densities. Convenience is an inherent advantage of DMS analysis owing to the minimal pre-processing of samples and automated sampling. The maintenance costs of DMS are low and the technology is robust.

The cell counts detected in this study correspond to pre-clinical tumors. Human tumors need neoangiogenesis to grow past the size of 10⁶ cells and 2 mm in diameter. Tumors below this size are rarely significant unless they secrete hormones [34]. Mammography screening is the means for breast cancer detection at a pre-clinical stage [2]. To provide clinical context to tumor size, the sensitivity of mammography imaging for breast cancer detection increases from 0% to 85%, as the tumor size increases from 2 mm to 20 mm [35]. The highest cell density in our study was 37 000 cells μL^{-1} . These ranges give a perspective to what constitutes a clinically relevant number of cancer cells. Moreover, they highlight the fact that the detection threshold of DMS for breast cancer cells is sufficient for clinical purposes.

The cell densities analyzed in this study can also be compared with the number of cancer cells required to grow a tumor in a murine model. Syngeneic murine models are considered the best animal models for human breast cancer research [36]. They are produced by implanting mouse-derived breast cancer cells into immune competent mice. The number of cells injected to induce tumor growth typically vary between 10⁴ and 10⁵ [37–42]. We detected cell densities of 3700 cells μL^{-1} with DMS. Of this density, approximately 6400 cells reached DMS analysis. As described earlier, this number of cells would not be sufficient in growing a tumor in most syngeneic murine models.

Most previous cell line detection studies have focused on using MS as the analysis method. Gas-chromatography (GC) and MS combined with solid-phase microextraction have been used to

analyze the volatile organic compound (VOC) profiles of breast cancer cell lines. Several VOC biomarkers indicative of breast cancer have been found, and benign and malignant cell lines can be differentiated according to their VOC profiles [43,44]. However, VOC extraction from cell cultures and GC are time-consuming and require several steps before MS.

Matrix-assisted laser desorption/ionization time-of-flight MS (MALDI-TOF MS) has been successfully applied in the characterization of various mammalian cell lines [45–48]. These studies have applied cell numbers equal to or larger than the two highest densities in our study. In general, the matrix used in MALDI-TOF MS is a customized α -Cyano-4-hydroxycinnamic acid (HCCA). In contrast, laser-based DMS does not require a specific matrix. The molecular composition of Myogel is more complex than that of HCCA, which results in a more complex background. Moreover, the duration of the MALDI-TOF MS sample preparation process varied between 2 min and more than 3 h [45,46], while DMS analysis requires minimal preparation. MALDI-TOF MS is a useful method for cell line identification, but it is not suitable for *ex vivo* tissue sample analysis, whereas DMS could be applied for both purposes.

In a recent study, Abu-Rabie et al. [49] studied rapid evaporative ionization MS (REIMS) with electrocautery sampling for discriminating TF1a and Jurkat cell lines (6.7×10^6 cells mL^{-1}). They achieved a discrimination rate between 94% and 100% in various MS configurations with leave-one-cell line-out CV. Contrary to the CV methods in our study, this CV does not take into consideration the potential drift in the instrumentation. The cell count in the Abu-Rabie et al. study was double the highest cell count in our study. This indicates the potential of DMS to provide a more cost-effective technology with a similar discriminatory capability as the REIMS.

Matrices derived from the mouse Engelbreth-Holm-Swarm (EHS) sarcoma, such as Matrigel, are widely used in *in vitro* cancer studies. However, the composition of all mouse tumor tissue homogenates differ from human TME. Myogel is a homogenous product similar to Matrigel, but it is extracted from human uterus leiomyoma tissue [20,50]. Thus, its protein composition differs significantly from the mouse EHS sarcoma derived products. However, both matrices contain basic extracellular matrix components, such as laminin, collagen IV, nidogen and EGF [20]. Myogel induces migration and invasion of carcinoma cells better than Matrigel [50–52]. Also, Myogel is reported to affect drug response in head and neck cancer cell lines significantly more than Matrigel [53]. Tuomainen et al. [53] reported that the cells cultured on Myogel reflect the response rates in clinical trials better than

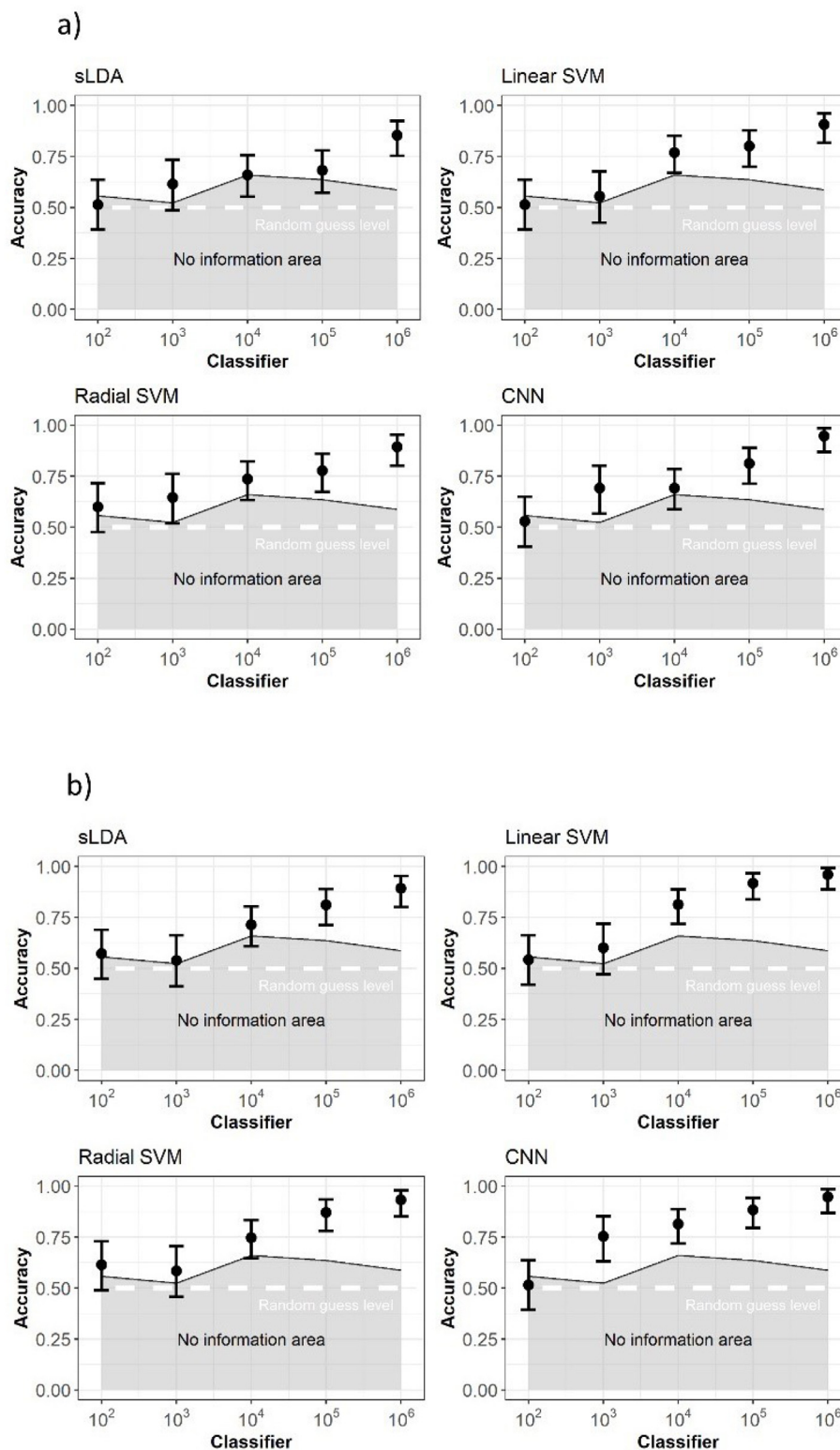


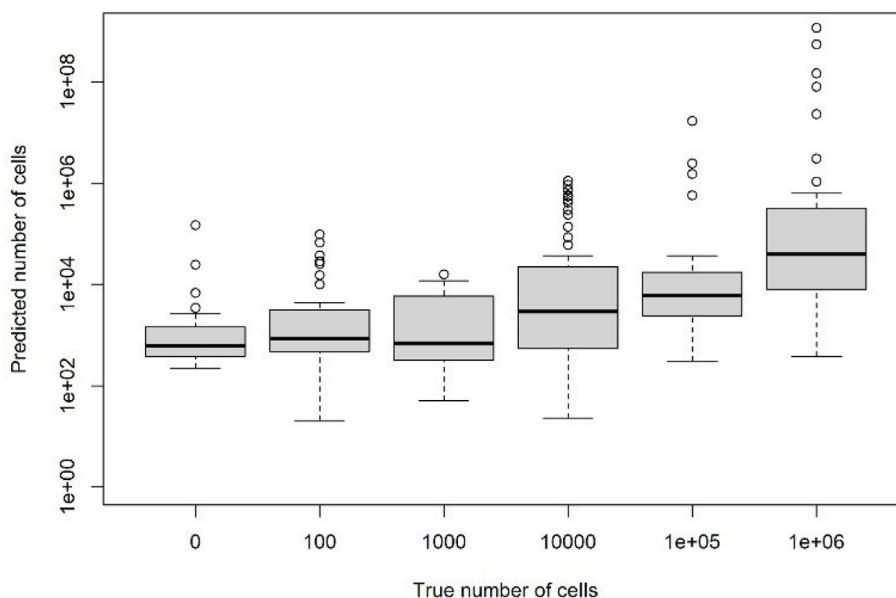
Fig. 4. The classification accuracies obtained with each classification model using A) day-wise cross-validation (CV) and B) well-wise CV. Each classifier had two classes: the cell density presented on the horizontal axis and the zero level. The colored area represents the guess level (maximum class). The error bars show the lower and upper limits for the accuracy estimates. The no-information rate and random guess level (50%) are presented to ease comparability of results.

Matrigel. Our aim was to ensure a reproducible and reliable study design by using a well-established breast cancer cell line and a controlled, homogenous matrix derived from human tumor tissue. These factors help overcome the issue of heterogeneity in human

breast cancer tumors obtained from surgical specimens.

The assessment of malignancy has been traditionally based on microscopic examination of cellular and architectural features of the tissue [54]. In case of cytology, only cellular features are

a) Day-wise CV



b) Well-wise CV

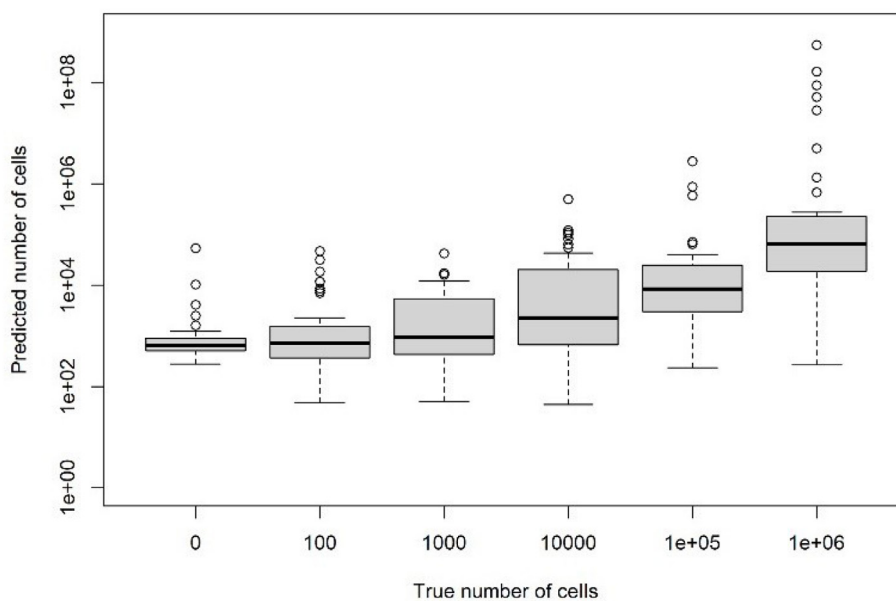


Fig. 5. The partial least squares regression results of the predicted cell densities as a boxplot according to A) day-wise cross-validation (CV) and B) well-wise CV.

available for examination [55]. These methods benefit from fixation and staining that typically takes days, which makes them unfeasible for applications that require rapid turnaround times, such as intraoperative margin evaluation. Frozen section and imprint cytology are specialized techniques that have been developed to enable rapid turnaround times of 30 min or less. Still, these

methods rely on a subjective interpretation of an expert pathologist [56]. Molecular assessment by DMS enables a rapid and an operator-independent way to assess malignancy.

We have previously applied CNNs and PLS regression in the prediction of phospholipid concentrations from DMS spectra and demonstrated that biologically relevant changes in lecithin

concentrations were detectable [33]. This is of interest as differences in the phospholipid compositions of smoke samples from healthy and cancerous tissues largely account for their MS identification [12]. In this study, we applied a more complex study design and were able to identify very low cancer cell densities from a cancer cell-free Myogel matrix mixture. These results validate DMS in the identification of densities relevant to margin detection and emphasize the feasibility of DMS in future clinical setups.

The limitations of the study should be taken into account. First, the purpose of this study was proof-of-concept. The number of independent samples was relatively small, which limits the generalizability of the results. The three lowest cell densities were not detectable with DMS, and more data is likely required to improve the detection of even the highest densities. Also, due to the prototype nature of the system, a small proportion of samples needed to be excluded due to the deficiencies in sample preparation and malfunctions in the DMS analysis. Second, the use of only one cell line does not comprehensively represent breast cancer, which is a heterogeneous disease. Last, although Myogel mimics human TME more precisely than Matrigel, it is also different from a real human breast cancer microenvironment limiting the generalizability of these results. Moreover, a larger study sample with even higher cell number ($\geq 10^7$) could have been beneficial in the assessment of detection levels.

Our future goal is to improve the components of the DMS device and provide faster and more accurate diagnostics at improved levels of detection. We aim to study the detection threshold in more complex and dynamic TMEs with multiple breast cancer cell lines to better mimic real-life cancer development. Human uterus leiomyoma-derived discs have previously been used in 3D cancer invasion assays [57], and they contain various cell types and matrix components commonly present in neoplastic stroma [50]. A study using myoma discs and several breast cancer cell lines could provide more information about the detection capability of DMS. In addition, a non-cancerous cell line could be used as a reference level to study the selectivity of the device.

5. Conclusions

The results demonstrate that a prototype DMS device combined with laser desorption is an effective method to distinguish low densities of breast cancer cells from Myogel *in vitro*. These results validate the use of DMS for identifying cell densities relevant to margin detection and emphasize the feasibility of DMS in future clinical setups. The results suggest that DMS analysis has great potential for intraoperative surgical margin assessment alongside MS-based methods. Moreover, laser-based DMS could be applied to support pathological analysis of *ex vivo* specimens and for quality control applications in biotechnology. This study design could act as a reproducible means to compare tissue analysis methods and their ability to discriminate between different densities of malignant and benign cells in the future.

CRedit authorship contribution statement

Lydia Lindfors: Investigation, Writing – original draft, Visualization, Writing – review & editing. **Patrik Sioris:** Investigation, Writing – original draft, Visualization, Writing – review & editing. **Anna Anttalainen:** Software, Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Katja Korelin:** Methodology, Writing – review & editing. **Anton Kontunen:** Software, Writing – original draft, Visualization, Writing – review & editing. **Markus Karjalainen:** Software, Writing – review & editing. **Erika Naakka:** Investigation, Resources, Writing – review & editing. **Tuula Salo:** Funding

acquisition, Supervision, Conceptualization, Methodology, Resources, Writing – review & editing. **Antti Vehkaoja:** Funding acquisition, Supervision, Resources, Writing – review & editing. **Niku Oksala:** Funding acquisition, Supervision, Resources, Conceptualization, Methodology, Writing – review & editing. **Vesa Hytönen:** Conceptualization, Methodology, Writing – review & editing. **Antti Roine:** Funding acquisition, Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing, Project administration. **Maiju Lepomäki:** Conceptualization, Methodology, Supervision, Project administration, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Antti Roine, Niku Oksala, Markus Karjalainen, and Anton Kontunen, are shareholders and employees of Olfactomics Ltd, a medical device company that develops novel technology for intraoperative surgical margin assessment. Anton Kontunen declares funding from the Doctoral School of Tampere University, The Finnish Foundation for Technology Promotion (Grant number 7671) and Emil Aaltonen Foundation (Grant number 210073). Markus Karjalainen declares funding from the Finnish Cultural Foundation, Pirkanmaa Regional Fund. Niku Oksala declares funding from Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital (Grant numbers 9s045, 9T044, 9U042, 150618, 9V044, 9X040, 9AA057, 9ab052 and Mk301; from Competitive funding to strengthen university research profiles funded by Academy of Finland, decision number 292377). Maiju Lepomäki declares funding from the Doctoral School of Tampere University, The Finnish Medical Foundation (grant numbers 2167, 4038), and Cancer Foundation of Finland. The Myogel research has been funded by Sigrid Jusélius Foundation, the Cancer Society of Finland, the Oulu University Hospital MRC grant, the Helsinki University Central Hospital research funds, and Jane and Aatos Erkkö Foundation. The Myogel research funding consisted of academic grants without any engagements considering the research project. The remaining authors declare that they have no conflicts of interest.

Acknowledgements

This study has received funding from the ATTRACT project funded by the EC under Grant Agreement 777222. The authors would like to thank Merja Jokela for assistance with the sample preparation, Artturi Vuorinen for participation in the ATLAS measurements and Meri Mäkelä for the proofreading of this paper.

References

- [1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries *CA Cancer, J. Clin.* 71 (2021) 209–249, <https://doi.org/10.3322/caac.21660>.
- [2] F. Cardoso, S. Kyriakides, S. Ohno, F. Penault-Llorca, P. Poortmans, I.T. Rubio, S. Zackrisson, E. Senkus, Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up, *Ann. Oncol.* 30 (2019) 1194, <https://doi.org/10.1093/annonc/mdz173>.
- [3] J. Fajdic, D. Djurovic, N. Gotovac, Z. Hrgovic, Criteria and procedures for breast conserving surgery, *Acta Inf. Med.* 21 (2013) 16–19, <https://doi.org/10.5455/aim.2013.21.16-19>.
- [4] E. Senkus, S. Kyriakides, S. Ohno, F. Penault-Llorca, P. Poortmans, E. Rutgers, S. Zackrisson, F. Cardoso, Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up, *Ann. Oncol.* 26 (2015) v8–v30, <https://doi.org/10.1093/annonc/mdv298>.
- [5] 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, M. Morrow, Society of Surgical Oncology-American Society for Radiation

- 10.1002/rcm.8525.
- [50] T. Salo, M. Dourado, E. Sundquist, E. Apu, I. Alahuhta, K. Tuomainen, J. Vasara, A. Al-Samadi, Organotypic three-dimensional assays based on human leiomyoma-derived matrices, *Philos. Trans. R. Soc. B-Biol. Sci.* 373 (2018), <https://doi.org/10.1098/rstb.2016.0482>.
- [51] W. Wahbi, E. Naakka, K. Tuomainen, I. Suleymanova, A. Arpalahti, I. Miinalainen, J. Vaananen, R. Grenman, O. Monni, A. Al-Samadi, T. Salo, The critical effects of matrices on cultured carcinoma cells: human tumor-derived matrix promotes cell invasive properties *Exp. Cell Res.* 389 (2020), 111885, <https://doi.org/10.1016/j.yexcr.2020.111885>.
- [52] E. Naakka, K. Tuomainen, H. Wistrand, M. Palkama, I. Suleymanova, A. Al-Samadi, T. Salo, fully human tumor-based matrix in three-dimensional spheroid invasion assay, *J. Vis. Exp.* (2019), <https://doi.org/10.3791/59567>.
- [53] K. Tuomainen, A. Al-Samadi, S. Potdar, L. Turunen, M. Turunen, P. Karhemo, P. Bergman, M. Risteli, P. Åström, R. Tiikkaja, R. Grenman, K. Wennerberg, O. Monni, T. Salo, Human tumor-derived matrix improves the predictability of head and neck cancer drug testing cancers, *Basel* 12 (2019) 92, <https://doi.org/10.3390/cancers12010092>.
- [54] T.T. Brunyé, E. Mercan, D.L. Weaver, J.G. Elmore, Accuracy is in the eyes of the pathologist: the visual interpretive process and diagnostic accuracy with digital whole slide images, *J. Biomed. Inf.* 66 (2017) 171–179, <https://doi.org/10.1016/j.jbi.2017.01.004>.
- [55] M.A. Al-Abbadi, Basics of cytology avicenna, *J. Med.* 1 (2011) 18–28, <https://doi.org/10.4103/2231-0770.83719>.
- [56] B.W. Maloney, D.M. McClatchy, B.W. Pogue, K.D. Paulsen, W.A. Wells, R.J. Barth, Review of methods for intraoperative margin detection for breast conserving surgery, *J. Biomed. Opt.* 23 (2018) 1–19, <https://doi.org/10.1117/1.JBO.23.10.100901>.
- [57] H. Tuominen, A. Al-Samadi, T. Salo, J. Rautava, Human myoma tissue-based extracellular matrix models for testing the effects of irradiation on the HPV positive cells *Virology*, 17 (2020) 87, <https://doi.org/10.1186/s12985-020-01367-1>.