

SCIENTIFIC REPORTS



OPEN

Label-free quantitative evaluation of breast tissue using Spatial Light Interference Microscopy (SLIM)

Hassan Majeed¹, Tan Huu Nguyen¹, Mikhail Eugene Kandel¹, Andre Kajdacsy-Balla² & Gabriel Popescu¹

Breast cancer is the most common type of cancer among women worldwide. The standard histopathology of breast tissue, the primary means of disease diagnosis, involves manual microscopic examination of stained tissue by a pathologist. Because this method relies on *qualitative* information, it can result in inter-observer variation. Furthermore, for difficult cases the pathologist often needs additional markers of malignancy to help in making a diagnosis, a need that can potentially be met by novel microscopy methods. We present a *quantitative* method for label-free breast tissue evaluation using Spatial Light Interference Microscopy (SLIM). By extracting tissue markers of malignancy based on the nanostructure revealed by the optical path-length, our method provides an objective, label-free and potentially automatable method for breast histopathology. We demonstrated our method by imaging a tissue microarray consisting of 68 different subjects — 34 with malignant and 34 with benign tissues. Three-fold cross validation results showed a sensitivity of 94% and specificity of 85% for detecting cancer. Our disease signatures represent intrinsic physical attributes of the sample, independent of staining quality, facilitating classification through machine learning packages since our images do not vary from scan to scan or instrument to instrument.

The latest World Health Organization (WHO) figures have reported breast cancer as the second most common form of cancer worldwide with 522,000 deaths in 2012¹. Within the US over 200,000 new cases of the disease are expected for women in 2017 according to the American Cancer Society². Effective treatment strategies require timely and accurate diagnosis of the disease. It has been reported that, in the US, the 5-year average survival rates for patients with invasive breast cancers increase from 90% to 99% when the disease is detected at a localized (non-metastatic) stage³.

The standard tissue evaluation method for diagnosing breast cancers involves microscopic examination of a hematoxylin and eosin (H&E) counter-stained tissue biopsy. The biopsy specimen is obtained from the patient when suspicion of disease is noted during a screening procedure such as X-ray mammography. Since cells and histological tissue sections are transparent, the H&E stain provides the necessary contrast for assessing tissue morphology using a conventional bright field microscope. This standard histopathology process has two important short-comings: reliance on qualitative markers leads to intra- and inter-observer variation while manual examination can lower the throughput of the evaluation. Furthermore, evaluation of traditional H&E based markers is sometimes not sufficient for stratifying difficult cases and additional markers are needed to provide further information^{4,5}. Novel quantitative microscopy methods can potentially help pathologists by offering an objective assessment of the tissue physical properties as well as new image markers for assessing difficult cases. Furthermore, quantitative markers can be interpreted by machine learning classifiers for rapid analysis and automated detection⁶.

In this work, we present a method for extracting quantitative markers of malignancy in breast tissue biopsies using Spatial Light Interference Microscopy (SLIM)⁷. SLIM is a quantitative phase imaging (QPI)⁸ modality that generates contrast by measuring the variation of optical path-length difference (OPD) across the tissue specimen. OPD reports on the product of the refractive index and thickness of tissue at each pixel. Malignant transformation involves physical changes in epithelial cell size and density as well as the tissue organization — both of which

¹Quantitative Light Imaging (QLI) Lab, Beckman Institute of Advanced Science and Technology, University of Illinois at Urbana Champaign, 405 N Matthews, Urbana, IL 61801, USA. ²Department of Pathology, University of Illinois at Chicago, 840 South Wood Street, Suite 130 CSN, Chicago, IL 60612, USA. Correspondence and requests for materials should be addressed to G.P. (email: gpopescu@illinois.edu)

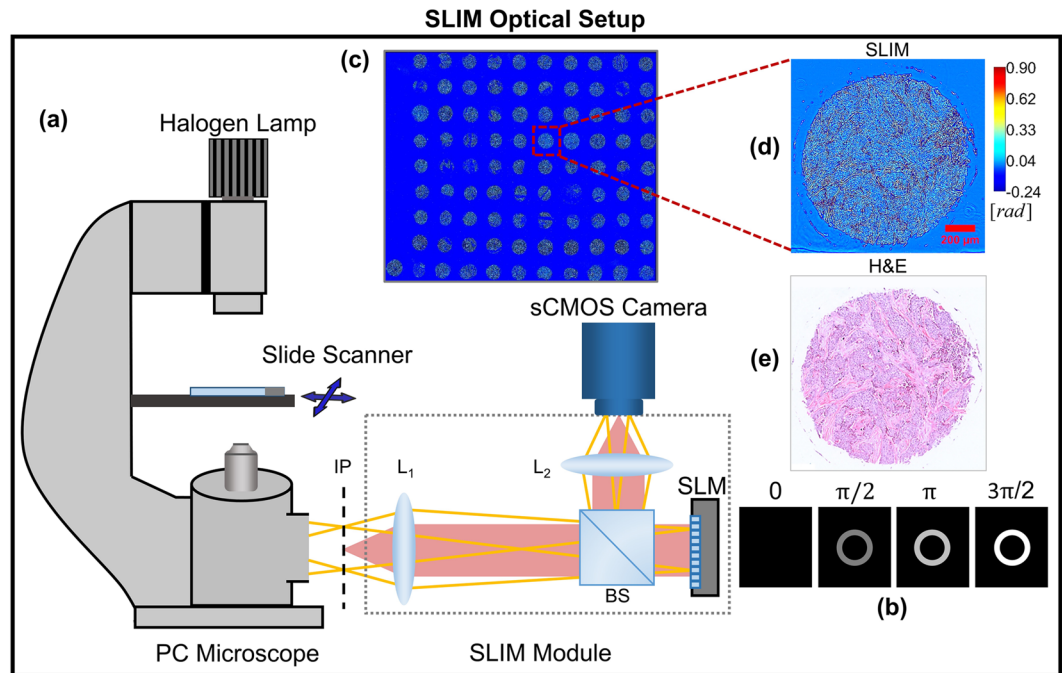


Figure 1. (a) The SLIM module added on to a commercial phase contrast microscope. (b) Four frames are acquired to compute one phase image by modulating the phase difference between scattered and incident light using a spatial light modulator (SLM). (c) An image of the whole slide scanned using SLIM. (d) Example of a TMA core SLIM image. (e) Bright field image of the same core after H&E staining. BS, beam splitter; L_1 - L_2 , lenses; IP, image plane.

affect OPD maps of tissue. Therefore, these maps have been used in the past for several clinical investigations⁹, including diagnosis of prostate¹⁰ and colorectal cancers¹¹, prediction of recurrence in prostate cancer¹², analysis of Gleason grade¹³, assessment of metastatic pancreatic cells¹⁴, as well as detection of pre-malignancy in colorectal tissue¹⁵. Furthermore, using QPI, other live human cells have also been investigated for morphological^{16,17}, chemical¹⁷⁻¹⁹ and mechanical markers of disease²⁰⁻²².

To date, the majority of quantitative image analyses on breast tissue biopsies have relied on color images of stained tissue. Image classification in these cases has involved computing a wide range of histological features including geometric features^{23,24}, texture-related features^{25,26} and radiometric features^{25,27,28} [see²⁹ for a review of methods]. However, the feature extraction process relies heavily on tissue staining which can vary from sample to sample and instrument to instrument, affecting the robustness of the classifier³⁰. The label-free approach we propose makes classification through machine learning easier since the instrument does not require calibration for inconsistency in pixel values due to variations in staining, tissue changes caused by harsh solvents etc. These advantages of QPI have already been leveraged to develop supervised learning methods for classifying erythrocytes infected with *Plasmodium falciparum*³¹ and non-activated lymphocytes³². Other label-free quantitative methods for tissue image classification have been proposed in the literature, including Fourier transform infrared spectroscopy (FTIR)³³⁻³⁵, Raman spectroscopy³⁶⁻³⁸, optical coherence tomography (OCT)^{39,40}, and second-harmonic generation (SHG) imaging^{41,42}. However, these techniques differ from our QPI-based method in terms of speed, resolution, and compatibility with the current diagnostic pipeline.

In our previous work we demonstrated that SLIM captures sufficient tissue morphology to separate benign from malignant tissue via visual investigation by trained pathologists⁴³. In that work, two board certified pathologists diagnosed tissue microarray (TMA) cores as either benign or malignant by examining tissue morphology in SLIM phase images. 88% and 87% accuracy of diagnosis using SLIM was obtained for the pathologists, respectively⁴³. In this work, we demonstrate the quantitative analysis capabilities of our tissue evaluation system by imaging a TMA comprising 68 different cases (34 benign and 34 malignant). For each epithelial region within a tissue core, we extracted scattering, geometric, and texture-related markers of tissue malignancy from the SLIM maps (see Materials and Methods). A linear-discriminant analysis (LDA) classifier was trained to separate benign cases from malignant cases and three-fold cross validation was performed to measure the classification accuracy of the learned model^{44,45}. Using validation by the Receiver Operating Characteristic (ROC) curve analysis, our results revealed a sensitivity of 94% and specificity of 85%. Our results are the first demonstration, to our knowledge, of using OPD based tissue markers for detecting malignancy in breast tissue, label-free.

Materials and Methods

SLIM Optical Setup. Figure 1 illustrates the SLIM optical setup which has been discussed in detail in previous publications^{7,46}. The setup comprises a module (CellVista SLIM Pro, Phi Optics, Inc.) coupled to the output port of a commercial phase contrast microscope (Carl Zeiss, Axio Observer Z1). This compatibility with existing

microscopes promises to reduce barriers to clinical adoption since optical microscopes are commonly available in pathology labs. In the SLIM module, the conjugate image plane outside the microscope is relayed onto a sCMOS camera (Andor, Zyla) using a 4f system comprising lenses L_1 and L_2 . At the Fourier plane of L_1 , a spatial light modulator (Boulder Nonlinear Systems) is used to modulate the phase difference between the scattered and unscattered components of light in increments of $\pi/2$. Four different modulations are applied [Fig. 1(b)] and the resulting phase image is reconstructed using a previously published algorithm⁷. Using a software platform developed in-house, the SLIM module has been upgraded with full-slide scanning capabilities^{11,43}. The acquisition speed is in the range of the existing commercial tissue scanners, which, in turn, only perform bright field imaging^{11,43}. Throughout our experiments, a 40x/0.75 NA phase contrast objective was used for imaging. At this sampling rate (6.2 pixels/ μm), the typical time for imaging a single tissue core (1 mm² area) was approximately 12 sec.

Tissue microarray. The TMA used for our study was purchased from US Biomax Inc. (Serial # BR-1002) with diagnosis for each case provided by the manufacturer through examination by a board certified pathologist. The TMA was obtained with all human subject information de-identified. Neither the authors of this work nor their institutions were involved in the collection of tissue. The TMA consisted of cores 1 mm in diameter and a section thickness of 5 μm . Standard formalin fixation and paraffin embedding (FFPE) histological preparation was used for each tissue block before extraction of cores. A xylene based mounting medium was used during cover-slipping.

The TMA consisted of 36 cases of infiltrating ductal carcinoma (IDC), 36 cases of tumor adjacent normal tissue and 10 cases of normal breast tissue derived from autopsy procedures (one core per case). Three of the tumor adjacent normal cores were obtained from the IDC cohort. The TMA was designed to mix approximately equal numbers of histologically normal and histologically invasive carcinoma cases. The post-mortem interval for autopsy cases was less than 6 hours and after assembly the TMA was inspected for quality control and diagnosis by the manufacturer through review by a board certified pathologist. For final analysis we selected 34 cores diagnosed as malignant and 34 cores diagnosed as normal (within which 28 were tumor adjacent normal and 6 were normal). Each of these final cores were selected based on whether the core was intact and whether any epithelial tissue was present in the core (cores containing only stromal tissue were excluded).

A SLIM image of the whole TMA slide is illustrated in Fig. 1(c). Figure 1(d,e) show, respectively, the phase map and the H&E stained tissue bright field image (henceforth referred to as ‘H&E image’) of one core. For obtaining a mosaic of the TMA, we used a C++ based stitching code, developed in-house¹¹. After staining the same tissue slide using standard protocols⁴⁷, H&E images of the TMA were acquired using a bright-field microscope (Carl Zeiss, Axio Observer Z1) outfitted with a color camera (Carl Zeiss, AxioCam MRC). The H&E images were used for qualitative evaluation only, to assist with annotation of epithelial regions in tissue, discussed below.

Annotation of epithelial regions in tissue images. Glands or continuous epithelial regions within each core were manually annotated using the region of interest (ROI) tool of ImageJ to allow feature extraction for each gland. A consistent criterion for annotation was used where groups of epithelial cells bounded by stroma on all sides were considered a single gland. Other tissue components within epithelium (such as lumen etc.) were considered part of the gland if bounded on all sides by epithelial cells. Glands from cores in the cancer cohort were labelled as malignant while those from cores in the normal cohort were labelled as benign.

Extraction of geometric and scattering features. Malignant transformation in breast tissue affects the size, shape and density of epithelial cells as well as the shape and organization of epithelial tissue. As a result, both the geometry and scattering properties of the gland are affected. We used gland perimeter curvature C , as well as the mean scattering length l_s , as part of the feature set used for separating benign and malignant tissue. The parameter extraction process is illustrated in Fig. 2 and a detailed description for each is provided below.

The extrinsic curvature C of a two-dimensional plane curve $P(x, y)$, that is parametrized by Cartesian coordinates $x(t)$ and $y(t)$ with parameter t , is given by the expression⁴⁸

$$C(t) = \frac{|x'y'' - y'x''|}{(x'^2 + y'^2)^{\frac{3}{2}}}, \quad (1)$$

where the x' , y' and x'' , y'' refer to the first and second derivatives in t , respectively. In the above parametrization, t refers to each pixel comprising the curve $P(x, y)$, having coordinates $x(t)$ and $y(t)$. This curvature can be interpreted as the magnitude of the rate of change of a vector tangent to $P(x, y)$. We computed C for the perimeter $P(x, y)$ of each annotated gland by using an open source MATLAB code⁴⁹. The code approximates $P(x, y)$ as a polygon before computing C for each point defining the gland perimeter, as described in Eq. (1). To speed up computation, the image of each core was first down-sampled from the raw image size of 8000 \times 8000 to 2048 \times 2048 pixels. The sampling rate in the down-sampled image was 1.59 pixels/ μm and a bi-cubic interpolation technique was used for down-sampling. The perimeter $P(x, y)$ was then further down-sampled by a factor of 20 (every 20th pixel was analyzed) before computing $C(t)$ in order to remove any pixel level errors due to manual annotation. The median gland curvature $\langle C \rangle$ was then used as a feature for separating benign and malignant cases. Figure 2(c) and (d) illustrate $\langle C \rangle$ for representative benign and malignant glands.

The scattering mean free path l_s , is a bulk scattering parameter that defines the length scale over which a single scattering event occurs on average. Assuming that the tissue slice captures the refractive index spatial fluctuation statistics, i.e., assuming statistical homogeneity, l_s can be computed through the *scattering-phase theorem* using the expression⁵⁰

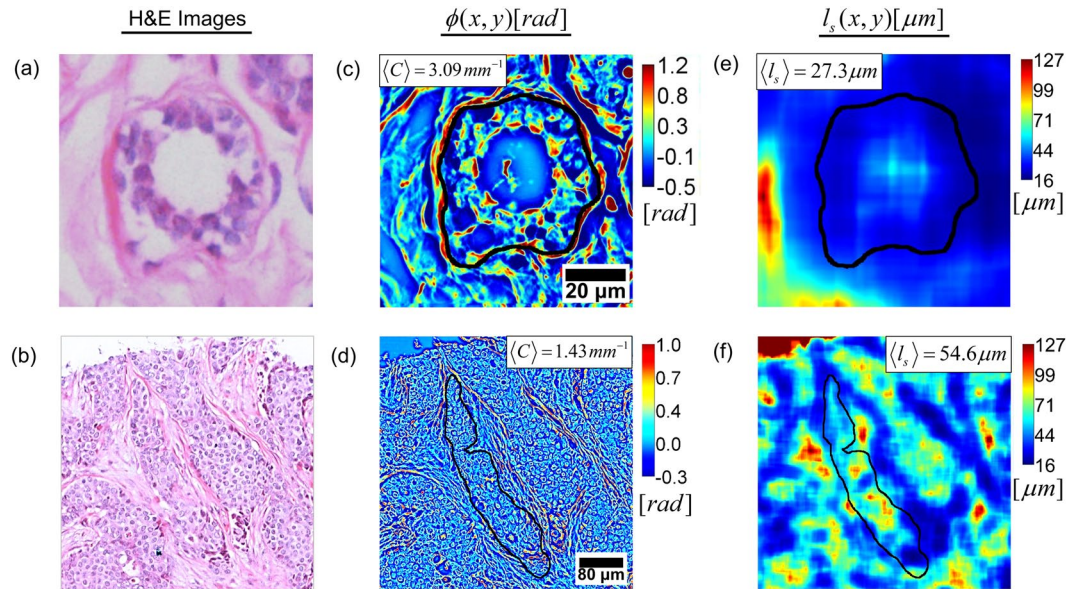


Figure 2. Computing the geometric feature (C) and scattering feature (l_s) over each annotated gland. (a) and (b) H&E images of benign and malignant glands, respectively. (c) and (d) SLIM images of the same benign and malignant glands, respectively, illustrating gland curvature C . The median over gland $\langle C \rangle$ is used as the geometric feature for classification. (e) and (f) $l_s(x, y)$ for benign and malignant glands, respectively. The median over gland $\langle l_s \rangle$ is used as the scattering feature for classification.

$$l_s = \frac{L}{\text{var}[\phi(x, y)]}, \quad (2)$$

where $\phi(x, y)$ is the SLIM phase image, L is the tissue section thickness and the operator $\text{var}[\cdot]$ computes the spatial variance over a region. The l_s parameter has been used in the past for discriminating between benign and malignant prostate tissue¹⁰. We first computed the image $l_s(x, y)$ from the phase image $\phi(x, y)$ (8000×8000 pixels) using a variance filter kernel size of 149×149 pixels, which equals the approximate diameter of 3 epithelial cells. The feature $\langle l_s \rangle$ was then computed by calculating the median of $l_s(x, y)$ over the gland area. This computation is illustrated in Fig. 2(e,f).

Extraction of texture-related features. Benign and malignant epithelial tissues differ not only in cell morphology but also in the organization of their components, leading to different textures. Texture-related features have been used in the past for solving different classification problems in histopathology of cancers^{13,29}. Our feature extraction follows the work done by Varma *et al.*⁵¹ for classifying different materials based on their texture. The approach is illustrated in Fig. 3. Each TMA core phase image was first down sampled to 2048×2048 pixels from 8000×8000 pixels [Fig. 3(a)]. A bi-cubic interpolation was used for down-sampling and the sampling rate in the down-sampled image was $1.59 \text{ pixels}/\mu\text{m}$. The down-sampled core image was then filtered through a convolution with the Leung-Malik (LM) filter bank [Fig. 3(b)]. This filter bank consists of gradient filters (both odd and even) at different orientations and spatial scales⁵². In total, 58 different filters were used, generating a 58-dimensional response vector for each pixel in the core phase image [Fig. 3(c)]. The response vectors from epithelial regions within each core were then randomly sampled (10000 vectors per core) to generate a smaller dataset for further processing. K-means clustering was then performed on the response vectors (number of clusters, $K = 50$) sampled from all cores within each training set (see Results and Discussion) and the computed cluster centroids were referred to as ‘textons’^{51,52}. $K = 50$ was chosen iteratively by repeatedly measuring the cross-validation AUC (see Results and Discussion) and determining the number of clusters required to maximize it (to account for both overfitting and separation accuracy). Since each pixel in each core belongs to a texton, for each pixel the histogram of textons was generated for its vicinity (window size 60×60 pixels) and was used to characterize the local texture in that neighborhood. This way, a 50 dimensional feature vector T was generated to characterize texture in a pixel’s neighborhood. An open source MATLAB code was used for generating the LM filter bank for this work⁵³.

Classifier training and validation. Since our work involves classifying each gland within a tissue core as benign or malignant, a feature vector for each gland was next generated by concatenating geometric, scattering and texture-related features. This procedure is illustrated in Fig. 4. After pixel-wise computation of gland curvature C , scattering length l_s and texture vector T , the median of each feature was computed over each gland in a core and a combined 52 dimension feature vector was generated for training. For each gland, this feature vector was then used as a predictor for training a linear-discriminant analysis (LDA) classifier [Fig. 4(a)]. Class labels,

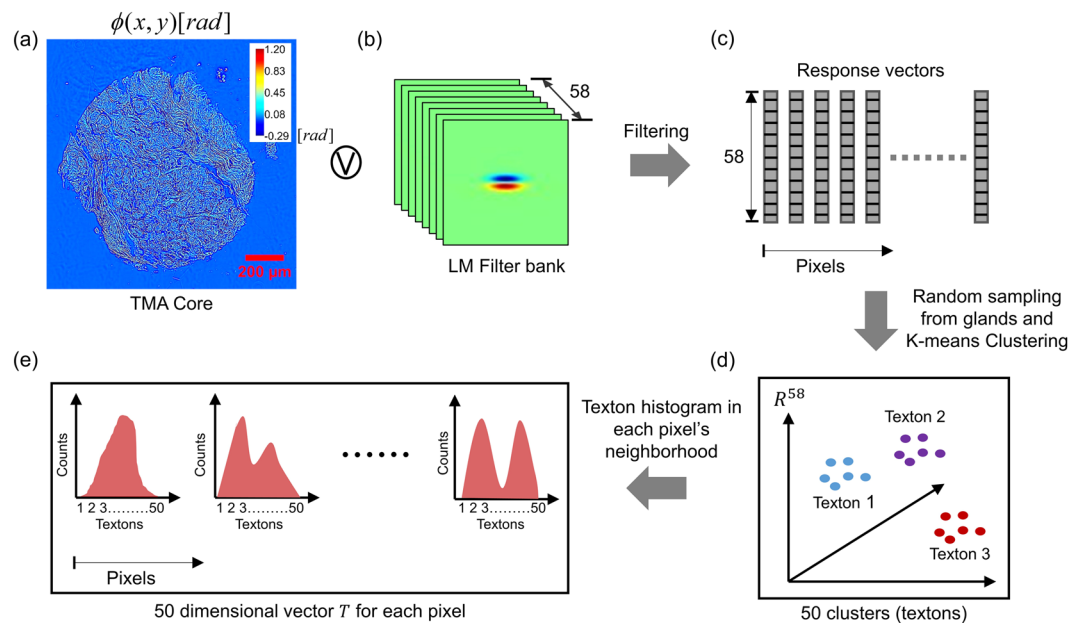


Figure 3. Algorithm for computing the texture in a pixel's neighborhood. (a–c) Generating the response of each pixel to a Leung-Malik filter bank. (d) K-means clustering of response vectors, generated from all cores in the training set, in order to find 50 cluster centroids or textons. (e) Histogram of textons, within a pixel's neighborhood, comprise the texture-related feature vector T for each pixel.

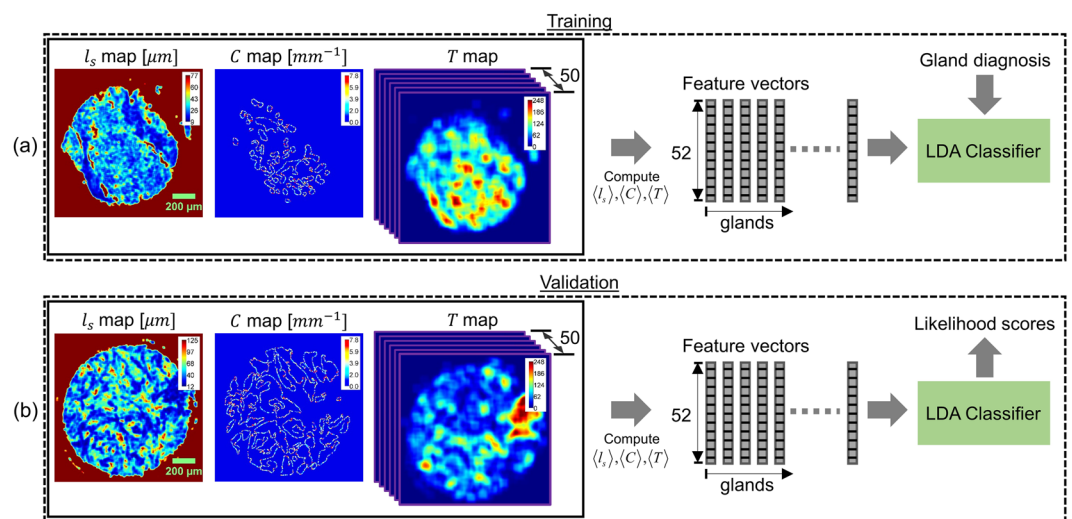


Figure 4. (a) Training and (b) Validation procedure for classifying glands as benign or malignant.

either benign or malignant, were used as the ground-truth for each gland during the training process. All glands within cores deemed cancerous by the pathologist were labelled malignant and all glands within cores deemed normal were labelled benign.

The feature extraction for validation purposes, illustrated in Fig. 4(b), followed a nearly identical procedure to that used during training. The only difference was that, instead of finding new textons (cluster centroids) for validation data, the texture feature vector T was computed by using the same textons as determined during training. As in training, a 52 dimensional feature vector was input to the LDA classifier which then used the model learned during training to generate a likelihood score for a gland being benign or malignant. Finally, the mean of the likelihood scores of all glands within a core was computed and used as the likelihood score of a core being benign or malignant. These scores were then used to generate a receiver operative characteristic (ROC) to select an operating point for separating benign and malignant cases (see Results and Discussion). For an annotated test core the total time required by our algorithm to generate a core likelihood score was approximately 2 minutes. This is in the absence of any parallelization of the computation through graphics processing unit (GPU) implementation which can significantly boost the computational throughput.

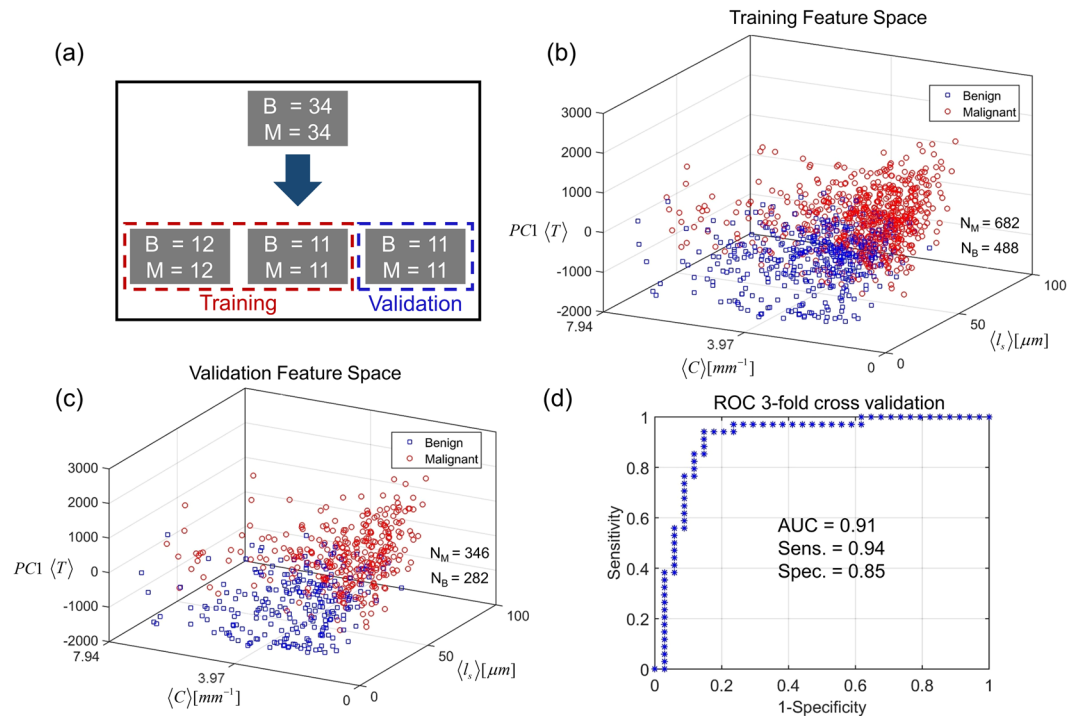


Figure 5. (a) Three-fold cross-validation procedure for evaluating classification accuracy. (b) Separation of benign and malignant gland feature vectors during *training* in 1 of 3 validation trials. (c) Separation of benign and malignant gland feature vectors during *validation* in 1 of 3 validation trials. (d) ROC curve for the 3 validation trials resulting in a sensitivity of 0.94 and specificity of 0.85 at the optimum operating point.

Data availability statement. The datasets generated or analyzed during this work are available upon reasonable request.

Results and Discussion

The classification results of our analysis are summarized in Fig. 5. In order to evaluate the accuracy of our method, we performed three-fold cross-validation⁵⁴ as illustrated in Fig. 5(a). The total number of cases were divided into three (nearly) equal groups. In each trial, two groups were used for training while the remaining one was used for validation. Thus, three validation trials were performed, each time selecting a different validation/training set combination.

Figure 5(b) illustrates the separation between benign and malignant gland feature vectors in one training set. In order to illustrate the data separation in 3 dimensions, we use principal component analysis (PCA) and represent the 50-dimensional feature vector $\langle T \rangle$ through its first principal component $PC1 \langle T \rangle$. The training space shows that scattering feature $\langle l_s \rangle$ has on average higher values for malignant glands than for benign glands. This finding is compatible with typical gland morphology in breast tissue since benign glands are well differentiated, consisting of a number of different structures including epithelial cells, lumen and myoepithelial cells⁵⁵. This heterogeneity of structure results in short mean scattering lengths as explained by a large variance in Eq. (2). Malignant glands on the other hand consist of a monoclonal proliferation of cells, sometimes even showing sheets of poorly differentiated epithelial cells, resulting in smaller variance and larger l_s values⁵⁵. These phenomena can also be observed in the examples given in Fig. 2(e,f). In previous investigations on prostate cancer, it was shown that l_s has a lower value in malignant tissue than in benign tissue¹⁰. That analysis, however, was carried out on larger areas of tissue where cellular organization can be different from the epithelial only regions we are studying in this work¹⁰.

The median gland curvature $\langle C \rangle$, on the other hand, generally has higher values for benign glands than for malignant glands. This is a result of the fact that the edge of a benign gland is constrained to follow a round or elliptical shape due to tubule formation [Fig. 2(a) and (c)]⁵⁵. When malignant transformation occurs, this constraint is broken and the gland edge is more irregular. At the spatial scale of investigation we have used here (approx. $13 \mu m$), the perimeter of the malignant gland is less rapidly varying, on average, than that of a benign gland. This geometric feature is similar to the previous measurement of the gland perimeter fractal dimension that has been used for histopathology^{23,29}.

Figure 5(c) shows the separation between benign and malignant glands in the validation feature space, where, qualitatively, the same separation trend is seen as in training. We show the results of only one of the three validation trials that were carried out. As described in Materials and Methods, the gland likelihood scores, generated by the classifier during validation, were averaged over each gland in order to obtain core-wise or case-wise scores. The standard deviation of gland likelihood scores, within each core, had a mean value of 0.18, a median value of 0.19 and a maximum value of 0.41 for the benign dataset (i.e. across 34 cores). The same for the malignant dataset

(again across 34 cores) were 0.17, 0.17 and 0.37, respectively. This indicates that the malignant cores consisted primarily (if not entirely) of malignant glands since the variance of gland likelihood scores over them was similar to that over benign cores, which consisted entirely of benign glands. In addition, while we treated all glands within cancerous cores as malignant during training, removing any benign tissue from malignant cores, during classifier training, is likely to improve our accuracy rather than worsen it.

The core-wise likelihood scores from the 3 trials were then pooled together to generate the ROC curve illustrated in Fig. 5(d)⁵⁶. Our results indicate an area under the curve (AUC) of 0.91. The optimum operating point for classification was determined by using the standard method of assigning equal weight to the cost of misclassifying positives and the cost of misclassifying negatives⁵⁷. This resulted in a sensitivity of 0.94 and specificity of 0.85 for the three-fold cross validation. When the same analysis was repeated by using the median rather than the mean of the gland scores over each core (to get a case-wise score), it resulted in an AUC of 0.91 and sensitivity and specificity of 0.94 and 0.82, respectively. While we use the standard method for determining the operating point here (which assigns equal cost to false negatives and false positives) in principle any operating point along the ROC curve can be chosen depending on the application (e.g. 0.97 sensitivity and 0.77 specificity). Having said that, the higher sensitivity of 0.94 is useful since it results in a smaller number of false negatives than false positives. Minimizing false negatives is more important than minimizing false positives since the latter only result in further investigations of the patient whereas the former constitute a missed diagnosis. A sensitive method is also useful in situations where a small biopsy specimen is available and detection of small amounts of malignant tissue visually is a challenge for the pathologist.

Our results are significant because they are the first illustration of tissue OPD derived features being used to detect intrinsic markers of malignancy in breast tissue. While previous works employing image analysis and supervised learning for detecting cancerous regions in H&E-stained tissue images have demonstrated good classifications AUCs (greater than 0.90)^{23,29,58,59}, accounting for stain variation through normalization remains a challenge³⁰ due to a lack of universal agreement on the correct normalization method²⁹. Our label-free results, thus, eliminate an important factor affecting consistency of results between different samples and instruments whilst maintaining high sensitivity and specificity.

Summary and Conclusions

In summary, we presented a new method for quantitative evaluation of tissue biopsies obtained from patients under investigation for breast cancer. Since our method relies on measurement of OPD maps, an intrinsic property of tissue, the basis for classification is objective and not subject to inter-observer variation. In the past much of quantitative histopathology has relied on analysis of stained tissue. However, stain variability continues to remain a grand challenge in applying computer algorithms across multiple H&E stained specimens. This fact is well documented in the literature, as follows. In ref.⁶⁰, it is stated that: “On the technical side, one of the main challenges in the computational interpretation of digital slide images has to do with color variations in the tissue induced by differences in slide preparation, staining, and even whole slide scanners. Clearly decision support algorithms that aim to work on digital pathology images will have to contend with and be resilient to these variations.” In ref.⁶¹ we find that “One of the major difficulties in breast cancer histopathology image analysis, particularly of H&E stained sections, is appearance variability.” As a final example, ref.⁶² states “It is clear that an integral part of digital pathology that has yet to be resolved is colour standardization; in order to do so, further work is needed focusing upon fine-tuning colour calibration methods in relation to the effect on diagnosis.”

Our method performs image processing and machine learning on unlabeled images, making it insensitive to variability due to staining. Although OPD depends on tissue slice thickness which can vary slightly from section to section, previous studies on colorectal and prostate cancer, using SLIM, have indicated that variations in OPD due to cutting errors are insignificant^{11,63}. These studies looked at the variation of median phase values and anisotropy in scattering between tissue slices having the same nominal thickness and reported insignificant differences. Despite this preliminary evidence, future studies where variation in our feature set is explicitly tested against tissue slice thickness variation are required. While in this work we have performed manual segmentation of glands, the automation of the entire process (including segmentation) is feasible and subject to future efforts.

Our results demonstrate, for the first time to our knowledge, that QPI based markers can be used for detecting malignancy in label-free breast histology samples at high accuracy. While our cross-validation results show promising sensitivity and specificity, a number of further studies are proposed before clinical adoption of our method. First, our analysis needs to be further tested with separate training and testing tests, the latter being obtained from an independent laboratory and remaining unused during model development. Second, our feature set needs to be applied to more clinically difficult cases since making frank benign versus frank malignant diagnoses is not currently a serious challenge for pathologists. However, our results demonstrate an important and necessary first step and we propose to apply our feature set, in future studies, to more challenging cases such as stratification of benign lesions as well as distinguishing ductal carcinoma *in situ* (DCIS) from ductal hyperplasia atypia^{4,5}. Third, while we have applied our diagnosis method to cases of IDC, the most widely prevalent form of breast cancer, the applicability of the model to other histological sub-types (such as infiltrating lobular carcinoma) also needs to be explored.

While other label-free diagnosis methods have been proposed for these types of investigations, they affect the standard diagnostic pipeline in terms of either speed, resolution or compatibility with established workflow. SLIM, on the other hand, requires minimal changes to a conventional microscopic optical train due to its modular design. Even though in this study a research grade microscope was coupled to the SLIM module, in principle SLIM can be used with any phase contrast microscope. Furthermore, the SLIM instrument is less expensive than most commercial tissue scanners. Our results, although preliminary, are an important stepping stone towards extracting reliable novel markers that can provide pathologists adjunct information to H&E based markers for assessing difficult cases. In addition, computational pathology tools such as ours can help pathologists where the biopsy specimen is small and highly sensitive detection is desired.

References

1. International Agency for Research on Cancer. *Cancer Fact Sheets: Breast Cancer*, <http://gco.iarc.fr/today/fact-sheets-cancers?cancer=15&type=0&sex=2> (2012).
2. American Cancer Society. *Cancer Facts & Figures*. (American Cancer Society, Atlanta, 2017).
3. American Cancer Society. *Cancer Prevention & Early Detection Facts & Figures 2017–2018*. (American Cancer Society, Atlanta, 2017).
4. Sloane, J. P. *et al.* Consistency of histopathological reporting of breast lesions detected by screening: findings of the UK. National External Quality Assessment (EQA) Scheme. U.K. National Coordinating Group for Breast Screening Pathology. *Eur. J. Cancer* **30A**, 1414–1419 (1994).
5. Elmore, J. G. *et al.* Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* **313**, 1122–1132 (2015).
6. Fernandez, D. C., Bhargava, R., Hewitt, S. M. & Levin, I. W. Infrared spectroscopic imaging for histopathologic recognition. *Nat. Biotechnol.* **23**, 469–474 (2005).
7. Wang, Z. *et al.* Spatial light interference microscopy (SLIM). *Opt. Express* **19**, 1016–1026 (2011).
8. Popescu, G. *Quantitative Phase Imaging of Cells and Tissues* (McGraw Hill, 2011).
9. Majeed, H. *et al.* Quantitative phase imaging for medical diagnosis. *J. Biophotonics* **10**, 177–205 (2017).
10. Wang, Z., Tangella, K., Balla, A. & Popescu, G. Tissue refractive index as marker of disease. *J. Biomed. Opt.* **16**, 116017 (2011).
11. Kandel, M. E. *et al.* Label-free tissue scanner for colorectal cancer screening. *J. Biomed. Opt.* **22**, 066016 (2017).
12. Sridharan, S., Macias, V., Tangella, K., Kajdacsy-Balla, A. & Popescu, G. Prediction of prostate cancer recurrence using quantitative phase imaging. *Sci. Rep.* **5** (2015).
13. Nguyen, T. H. *et al.* Automatic Gleason grading of prostate cancer using quantitative phase imaging and machine learning. *J. of Biomed. Opt.* **22**, 036015 (2017).
14. Kemper, B. *et al.* Investigation of living pancreas tumor cells by digital holographic microscopy. *J. Biomed. Opt.* **11**, 034005 (2006).
15. Uttam, S. *et al.* Early Prediction of Cancer Progression by Depth-Resolved Nanoscale Mapping of Nuclear Architecture from Unstained Tissue Specimens. *Cancer Res.* **75**, 4718–4727 (2015).
16. Mir, M., Wang, Z., Tangella, K. & Popescu, G. Diffraction Phase Cytometry: blood on a CD-ROM. *Opt. Express* **17**, 2579–2585 (2009).
17. Kim, Y. *et al.* Profiling individual human red blood cells using common-path diffraction optical tomography. *Sci. Rep.* **4**, 6659 (2014).
18. Park, Y. K. *et al.* Refractive index maps and membrane dynamics of human red blood cells parasitized by *Plasmodium falciparum*. *Proc. Natl. Acad. Sci.* **105**, 13730 (2008).
19. Yoon, J. *et al.* Label-free characterization of white blood cells by measuring 3D refractive index maps. *Biomed. Opt. Express* **6**, 3865–3875 (2015).
20. Bhaduri, B., Kandel, M., Brugnara, C., Tangella, K. & Popescu, G. Optical Assay of Erythrocyte Function in Banked Blood. *Sci. Rep.* **4**, 6211 (2014).
21. Shaked, N. T., Satterwhite, L. L., Telen, M. J., Truskey, G. A. & Wax, A. Quantitative microscopy and nanoscopy of sickle red blood cells performed by wide field digital interferometry. *J. Biomed. Opt.* **16**, 030506 (2011).
22. Eldridge, W. J., Steelman, Z. A., Loomis, B. & Wax, A. Optical Phase Measurements of Disorder Strength Link Microstructure to Cell Stiffness. *Biophys. J.* **112**, 692–702 (2017).
23. Chan, A. & Tuszynski, J. A. Automatic prediction of tumour malignancy in breast cancer with fractal dimension. *R. Soc. Open Sci.* **3**, 160558 (2016).
24. Dundar, M. M. *et al.* Computerized classification of intraductal breast lesions using histopathological images. *IEEE Trans. Biomed. Eng.* **58**, 1977–1984 (2011).
25. Kowal, M., Filipczuk, P., Obuchowicz, A., Korbicz, J. & Monczak, R. Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. *Comput. Biol. Med.* **43**, 1563–1572 (2013).
26. Weyn, B. *et al.* Automated breast tumor diagnosis and grading based on wavelet chromatin texture description. *Cytometry* **33**, 32–40 (1998).
27. Filipczuk, P., Kowal, M. & Obuchowicz, A. Fuzzy Clustering and Adaptive Thresholding Based Segmentation Method for Breast Cancer Diagnosis in *Computer Recognition Systems 4* (eds Robert Burduk, Marek Kurzyński, Michał Woźniak, & Andrzej Żołnierok) 613–622 (Springer Berlin Heidelberg, 2011).
28. J Janowczyk, A. & Madabhushi, A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J. Path. Inform.* **7**, 29 (2016).
29. Gurcan, M. N. *et al.* Histopathological Image Analysis: A Review. *IEEE Rev. Biomed. Eng.* **2**, 147–171 (2009).
30. Azar, J., Busch, C. & Carlbom, I. Histological stain evaluation for machine learning applications. *J. Path. Inform.* **4**, 11 (2013).
31. P Park, *et al.* A. Automated Detection of *P. falciparum* Using Machine Learning Algorithms with Quantitative Phase Images of Unstained Cells. *PLoS One* **11**, (2016).
32. Yoon, J. *et al.* Identification of non-activated lymphocytes using three-dimensional refractive index tomography and machine learning. *Sci. Rep.* **7**, 6654 (2017).
33. Benard, A. *et al.* Infrared imaging in breast cancer: automated tissue component recognition and spectral characterization of breast cancer cells as well as the tumor microenvironment. *Analyst* **139**, 1044–1056 (2014).
34. Walsh, M. J., Holton, S. E., Kajdacsy-Balla, A. & Bhargava, R. Attenuated total reflectance Fourier-transform infrared spectroscopic imaging for breast histopathology. *Vib. Spectrosc.* **60**, 23–28 (2012).
35. Malins, D. C. *et al.* The etiology and prediction of breast cancer. Fourier transform-infrared spectroscopy reveals progressive alterations in breast DNA leading to a cancer-like phenotype in a high proportion of normal women. *Cancer* **75**, 503–517 (1995).
36. Chowdary, M. V. P., Kumar, K. K., Kurien, J., Mathew, S. & Krishna, C. M. Discrimination of normal, benign, and malignant breast tissues by Raman spectroscopy. *Biopolymers* **83**, 556–569 (2006).
37. Haka, A. S. *et al.* Diagnosing breast cancer by using Raman spectroscopy. *Proc. Natl. Acad. Sci.* **102**, 12371–12376 (2005).
38. Frank, C., McCreery, R. & Redd, D. C. B. Raman Spectroscopy of Normal and Diseased Human Breast Tissues. *Anal. Chem.* **67**, 777–783 (1995).
39. Nguyen, F. T. *et al.* Intraoperative Evaluation of Breast Tumor Margins with Optical Coherence Tomography. *Cancer Res.* **69**, 8790–8796 (2009).
40. Zysk, A. M. & Boppart, S. A. Computational methods for analysis of human breast tumor tissue in optical coherence tomography images. *J. Biomed. Opt.* **11**, 054015 (2006).
41. Bredfeldt, J. S. *et al.* Automated quantification of aligned collagen for human breast carcinoma prognosis. *J. Path. Inform.* **5**, 28 (2014).
42. Campagnola, P. Second Harmonic Generation Imaging Microscopy: Applications to Diseases Diagnostics. *Anal. Chem.* **83**, 3224–3231 (2011).
43. Majeed, H. *et al.* Breast cancer diagnosis using spatial light interference microscopy. *J. Biomed. Opt.* **20**, 111210 (2015).
44. Fisher, R. A. The Use Of Multiple Measurements In Taxonomic Problems. *Ann. Eugen.* **7**, 179–188 (1936).
45. Heang-Ping, C. *et al.* Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space. *Phys. Med. Biol.* **40**, 857 (1995).
46. Kim, T. *et al.* White-light diffraction tomography of unlabelled live cells. *Nat. Photonics* **8**, 256–263 (2014).

47. Fischer, A. H., Jacobson, K. A., Rose, J. & Zeller, R. Hematoxylin and Eosin Staining of Tissue and Cell Sections. *Cold Spring Harb. Protoc.* **2008**, pdb.prot4986 (2008).
48. Kreyszig, E. *Differential Geometry* (The University of Toronto Press, 1991).
49. Kroon, D.-J. 2D Line Curvature and Normals, *Mathworks File Exchange*, <https://www.mathworks.com/matlabcentral/fileexchange/32696-2d-line-curvature-and-normals?focused=5199905&tab=function> (2011).
50. Wang, Z., Ding, H. & Popescu, G. Scattering-phase theorem. *Optics letters* **36**, 1215–1217 (2011).
51. Varma, M. & Zisserman, A. A Statistical Approach to Texture Classification from Single Images. *Int. J. Comput. Vis.* **62**, 61–81 (2005).
52. Leung, T. & Malik, J. Representing and recognizing the visual appearance of materials using three-dimensional textons. *Int J Comput Vision* **43**, 29–44 (2001).
53. Visual Geometry Group. Matlab script for generating our implementation of the L-M filter bank, <http://www.robots.ox.ac.uk/~vgg/research/texclass/code/makeLMfilters.m> (2017).
54. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, 2009).
55. Kumar, V., Abbas, A. K. & Aster, J. C. *Robbins and Cotran Pathologic Basis of Disease* (Elsevier/Saunders, 2014).
56. Japkowicz, N. & Shah, M. ROC Curves in *Evaluating Learning Algorithms: A Classification Perspective*, 148–153 (Cambridge University Press, 2011).
57. Japkowicz, N. & Shah, M. Isometrics in *Evaluating Learning Algorithms: A Classification Perspective* 119–124 (Cambridge University Press, 2011).
58. Doyle, S., Agner, S., Madabhushi, A., Feldman, M. & Tomaszewski, J. In 2008 5th *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 496–499 (2008).
59. Araújo, T. *et al.* Classification of breast cancer histology images using Convolutional Neural Networks. *PLoS ONE* **12**, e0177544 (2017).
60. Madabhushi, A. & Lee, G. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical Image Analysis* **33**, 170–175 (2016).
61. Veta, M., Pluim, J. P. W., Diest, P. J. V. & Viergever, M. A. Breast Cancer Histopathology Image Analysis: A Review. *IEEE Trans. on Biomed. Eng.* **61**, 1400–1411 (2014).
62. Clarke, E. L. & Treanor, D. Colour in digital pathology: a review. *Histopathology* **70**, 153–163 (2017).
63. Sridharan, S. *et al.* Prediction of prostate cancer recurrence using quantitative phase imaging: Validation on a general population. *Sci. Rep.* **6** (2016).

Acknowledgements

We would like to thank Kingsley Boateng for his help with the H&E staining procedure for the tissue biopsies. This work was supported by the National Science Foundation (CBET-1040461 MRI, CBET-0939511 STC, DBI 1450962 EAGER and IIP-1353368). H. M's PhD research is being sponsored by the Beckman Graduate Fellowship Program administered through the support of the Arnold and Mabel Beckman Foundation.

Author Contributions

H.M. imaged the tissue samples and performed the image processing and analysis. He also wrote the main manuscript text and prepared the figures. T. N helped with the image processing and analysis, while M. K developed the tissue scanner used for imaging the samples. A. B provided clinical input and G. P edited the manuscript and supervised the project.

Additional Information

Competing Interests: G.P. is President and CEO of Phi Optics, Inc., a company that develops quantitative phase imaging technologies.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018