

AD \_\_\_\_\_  
(Leave blank)

Award Number:  
W81XWH-10-1-0112

TITLE:  
Epithelial and Stromal Spectral Imaging for Rapid Surgical Margin  
Analysis

PRINCIPAL INVESTIGATOR:  
Ashley Laughney, B.S.

CONTRACTING ORGANIZATION:  
Trustees of Dartmouth College  
Hanover, NH 03755-1404

REPORT DATE:  
March 2011

TYPE OF REPORT:  
Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: (Check one)

X Approved for public release; distribution unlimited

Á  
ÚåæÁ{æ}bÊÁ~\*æ^æ~^bÁá^äd~ãÁàæ^äæ^&bÁ'~^/áæ^æäÁæ^Á/áæbÁãæ\*~ã\ÁáãæÁ/á~bæ  
~ãÁ/áæÁá|\á~ãÇbDÁá^äÁbá~|→äÁ^~\ÁâæÁ'~^b\ã|æäÁábÁá^Á~ààæ^æ→Áææ\*áã\↑æ^\  
~ãÁ/áæÁÑã↑]Á\*~bæ\æ~^ÊÁ\*~→' ]Á~ãÁãæ'æbæ~^Á|^æbbÁb~Áãæb&^á\æäÁâ]Á~\áæã  
ä~'|↑æ^/á\æ~^ÊÁ

# REPORT DOCUMENTATION PAGE

*Form Approved*  
*OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD-MM-YYYY)</b> 31-MAR-2011			<b>2. REPORT TYPE</b> Annual Summary		<b>3. DATES COVERED (From - To)</b> 01 MAR 2010 - 28 FEB 2011	
<b>4. TITLE AND SUBTITLE</b>  "Epithelial and Stromal Spectral Imaging for Rapid Surgical Margin Analysis"					<b>5a. CONTRACT NUMBER</b> W81XWH-10-1-0112	
					<b>5b. GRANT NUMBER</b>	
					<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b> Ashley Laughney, ashley.laughney@dartmouth.edu					<b>5d. PROJECT NUMBER</b>	
					<b>5e. TASK NUMBER</b>	
					<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Trustees of Dartmouth College Jill Mortali Office of Sponsored Projects Hanover, NH 03755-1404					<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> U.S. Army Medical Research and Materiel Command 620 Chandler Street Fort Detrick, MD 21702-5012						
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>  Approved for public release; distribution unlimited					<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
					<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>13. SUPPLEMENTARY NOTES</b>						
<b>14. ABSTRACT</b> A broadband spectroscopy platform was developed to image thick tissue samples at a resolution sensitive to the diagnostic gold standard, pathology. Tissue samples were imaged in a 1cm <sup>2</sup> field of view across a static beam using a motorized stage, permitting wide-field optical characterization of diagnostic pathology. The sampling spot size confined the volume of tissue probed to within a few transport pathlengths so that multiple-scattering effects were minimized and simple empirical models parameterized the spectra. A k-Nearest Neighbor (k-NN) classifier was trained using parameters extracted from the localized scattering spectrum, automating diagnosis of benign and malignant breast pathologies in situ with a sensitivity and specificity of 91% and 77% respectively. Performance of the classifier was validated in 67,000 spectra from 29 excised breast tissues. The work effectively characterized the spectral response of breast pathologies and automated classification of the tissue's spectral response according to a diagnosis. Clinically feasible data acquisition speeds were attained through development of a dark-field in situ scanning-beam spectroscopy platform.						
<b>15. SUBJECT TERMS</b> Breast conserving surgery, scattering spectroscopy, classification, optics, spectral imaging						
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>  UU	<b>18. NUMBER OF PAGES</b>  20	<b>19a. NAME OF RESPONSIBLE PERSON</b> USAMRMC	
<b>a. REPORT</b> U	<b>b. ABSTRACT</b> U	<b>c. THIS PAGE</b> U			<b>19b. TELEPHONE NUMBER (include area code)</b>	

## Table of Contents

	<u>Page</u>
Introduction.....	1
Body.....	2
Key Research Accomplishments.....	11
Reportable Outcomes.....	12
Conclusion.....	13
References.....	14

## Introduction

Breast conserving therapy (BCT), which includes local excision and radiation treatment to the breast, has been the standard of care for early breast cancers (Stage 0-II), since five major prospective, randomized trials demonstrated that the long-term survival after BCT is equivalent to that of radical mastectomy for most patients when surgical margins are clear of residual disease<sup>1-5</sup>. Of the 250,000 women annually diagnosed with breast cancer eligible for BCT, approximately 165,000-180,000 have conservative surgery<sup>6</sup>. Following breast conservation, the strongest risk factor for local recurrence and mortality is a positive resection margin (tumor cells on ink). Therefore, if a margin is positive or close, the patient is advised to undergo re-excision surgery to achieve clear margins<sup>7,8</sup>. Margin status is currently evaluated post-operatively by microscopic evaluation of pathology in small, representative pieces of tissue. While sampled tissues are adequate for assessing tumor type, grade and receptor status, they are insufficient for evaluating important prognostic factors like disease extent and multi-focality. Additionally, breast tissue is markedly heterogeneous, which makes distinguishing small foci of cancer within the spectrum of normal tissue challenging when using point-based probe measurements<sup>9-11</sup>. A broadband spectroscopy platform was developed to image thick tissue samples at a resolution sensitive to the diagnostic gold standard, pathology<sup>12</sup>. Tissue samples were imaged in a 1cm<sup>2</sup> field of view across a static beam using a motorized stage, permitting wide-field optical characterization of diagnostic pathology. The sampling spot size (100µm lateral resolution) confined the volume of tissue probed to within a few transport pathlengths so that multiple-scattering effects were minimized and simple empirical models parameterized the spectra. A k-Nearest Neighbor (k-NN) classifier was trained using parameters extracted from the localized scattering spectrum, automating diagnosis of benign and malignant breast pathologies in situ with a sensitivity and specificity of 91% and 77% respectively. Performance of the classifier was validated in 67,000 spectra from 29 excised breast tissues<sup>13</sup>. The work achieved in year one of this Department of Defense Pre-doctoral Traineeship Award effectively characterized the spectral response of breast pathologies and automated classification of the tissue's spectral response according to a diagnosis. Clinically feasible data acquisition speeds were attained through development of a dark-field in situ scanning-beam spectroscopy platform. Year two of the traineeship fellowship will assess the ability of the spectral imaging platform to provide immediate evaluation of involved surgical margins for the presence of residual cancer during breast-conserving surgery.

## Body

Task 1. Assemble and parameterize an extensive databank of scatter spectra from fresh breast tissue across clinically relevant diagnostic categories.

### Materials and Methods

#### *1a. Scatter spectroscopy of fresh breast tissue*

Fresh breast tissue specimens were imaged in a custom-built micro-sampling reflectance spectral imaging system<sup>14</sup>. This system employs a quasi-confocal illumination and detection (510-785nm) to constrain the overlapping illumination and detection spot sizes to within approximately one scattering distance in tissue (~100 μm in the visible). A complete description of this imaging system can be found in a previous study<sup>15</sup>. Sampling in this mesoscopic regime allows the use of simple empirical models to describe the light transport. For the short pathlengths involved and for typical values of absorption and scattering in tissue, the measured spectral response is proportional to the reduced scattering coefficient,  $\mu_s'$ . In regions where significant local absorption is encountered, a Beer's law type attenuation factor is used to correct for the effects of absorption<sup>15-18</sup>.

#### *1b. Empirical model of spectral scattering response*

An empirical approximation to Mie theory was used to describe the measured reflectance spectrum,  $R(\lambda)$ , from a volume-averaged region of tissue<sup>19</sup>. Additionally, a Beer's law attenuation factor corrected for the presence of significant local absorption by Hemoglobin (Eq. 2)

$$R(\lambda) = A\lambda^{-b} \exp^{-\Gamma[HbT]\{SO_2*\epsilon_{HbO_2}(\lambda)+(1-SO_2)*\epsilon_{Hb}(\lambda)\}} \quad \text{Eq. 2}$$

Parameters  $A$  and  $b$  are scattering amplitude and scattering power respectively. Both depend on the size and number density of scattering centers in the volume of interrogated tissue, thereby reflecting variations in breast tissue morphology<sup>20-22</sup>.  $\Gamma$  refers to the mean optical pathlength,  $[HbT]$  is the total hemoglobin concentration,  $SO_2$  is the oxygen saturation factor (ratio of oxygenated to total hemoglobin), and  $\epsilon_{HbO_2}$  and  $\epsilon_{Hb}$  refer to the molar extinction coefficients of these two chromophores respectively<sup>23</sup>. Oxygenated and deoxygenated hemoglobin were the dominant tissue chromophores encountered in the measured waveband. Measured reflectance spectra were fit to this model using a nonlinear least squares solver to obtain estimates of scattering amplitude and scattering power relative to Spectralon. A measure of average scattering irradiance,  $I_{avg}$ , was calculated by integrating the reflectance spectrum over a waveband that avoids the hemoglobin absorption peaks (620:785nm). Scattering parameters were then microscopically correlated to morphological features identified by pathologist (Wendy Wells) on Hematoxylin and Eosin (H&E) stained sections of the tissue, cut in the exact same geometry as imaged in situ.

#### *1c. Associate scattering parameters with diagnostic pathology*

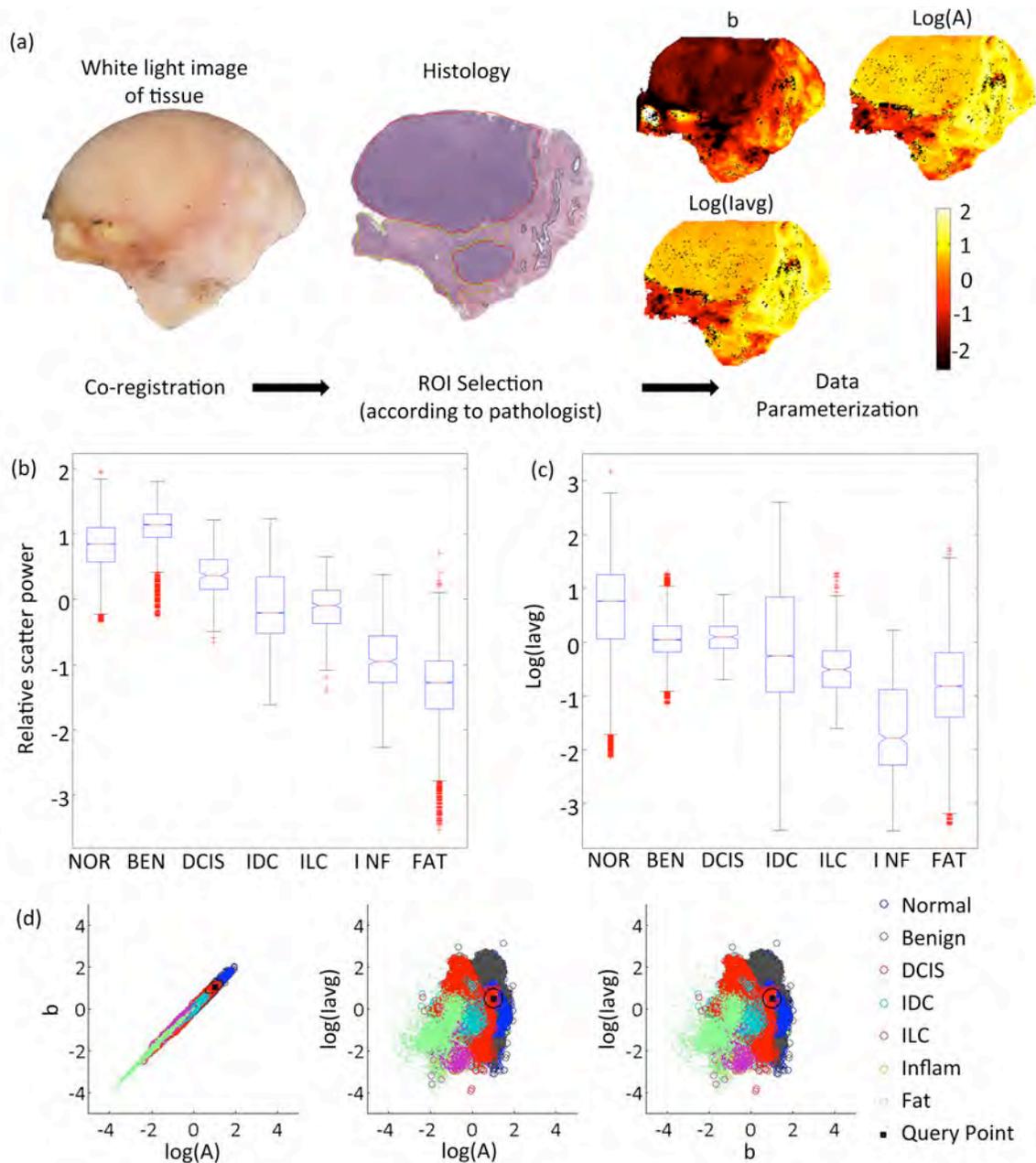
All studies were completed based upon a protocol approved by the Committee for the Protection of Human Subjects, Institutional Review Board (IRB) at Dartmouth. Fresh breast tissue was obtained directly from the Department of Pathology at Dartmouth-Hitchcock Medical Center from patients who had given informed consent to allow this use of their tissue. Samples were procured during conservative surgery or breast reduction surgery, and were only provided if there was tissue in excess of that required to make a clinical diagnosis. Tissue samples were 1-2 cm<sup>2</sup> with a thickness of ~3-5mm. Samples were imaged within 12 hours of surgery, and in the case of delay, the tissue was stored in a 4°C refrigerator and hydrated with a phosphate buffer solution. Immediately following imaging, each sample was placed in 10% formalin and processed for histology (paraffin embedded, sectioned to 4μm, and stained with H&E). A total of 35 tissue specimens were imaged; 6 were rejected from analysis due to low signal-to-noise and/or poor histological processing (both a consequence of highly fatty tissue). In the remaining 29 tissue samples, 48 regions of interest were identified by the pathologist and these are summarized in

Table 1. The pathologist identified seven tissue pathologies in the samples and these were classified more generally as not-malignant, malignant, or adipose.

<b>Tissue Type and Subtype</b>	<b># ROI</b>	<b># Spectra</b>
<b>Total Not-Malignant</b>	<b>25</b>	<b>36979</b>
<i>Normal Epithelium and Stroma</i>	21	31226
<i>Benign Epithelium and Stroma</i>	3	5220
<i>Inflammation</i>	1	533
<b>Total Malignant</b>	<b>14</b>	<b>23220</b>
<i>Ductal Carcinoma In Situ (DCIS)</i>	1	194
<i>Invasive Ductal Carcinoma (IDC)</i>	12	22547
<i>Invasive Lobular Carcinoma (ILC)</i>	1	479
<b>Total Adipose</b>	<b>9</b>	<b>7021</b>
<i>Adipose</i>	9	7021
<b>Total ROI</b>	<b>48</b>	<b>67220</b>

**Table 1 Distribution of the sample population according to tissue type and subtype.**

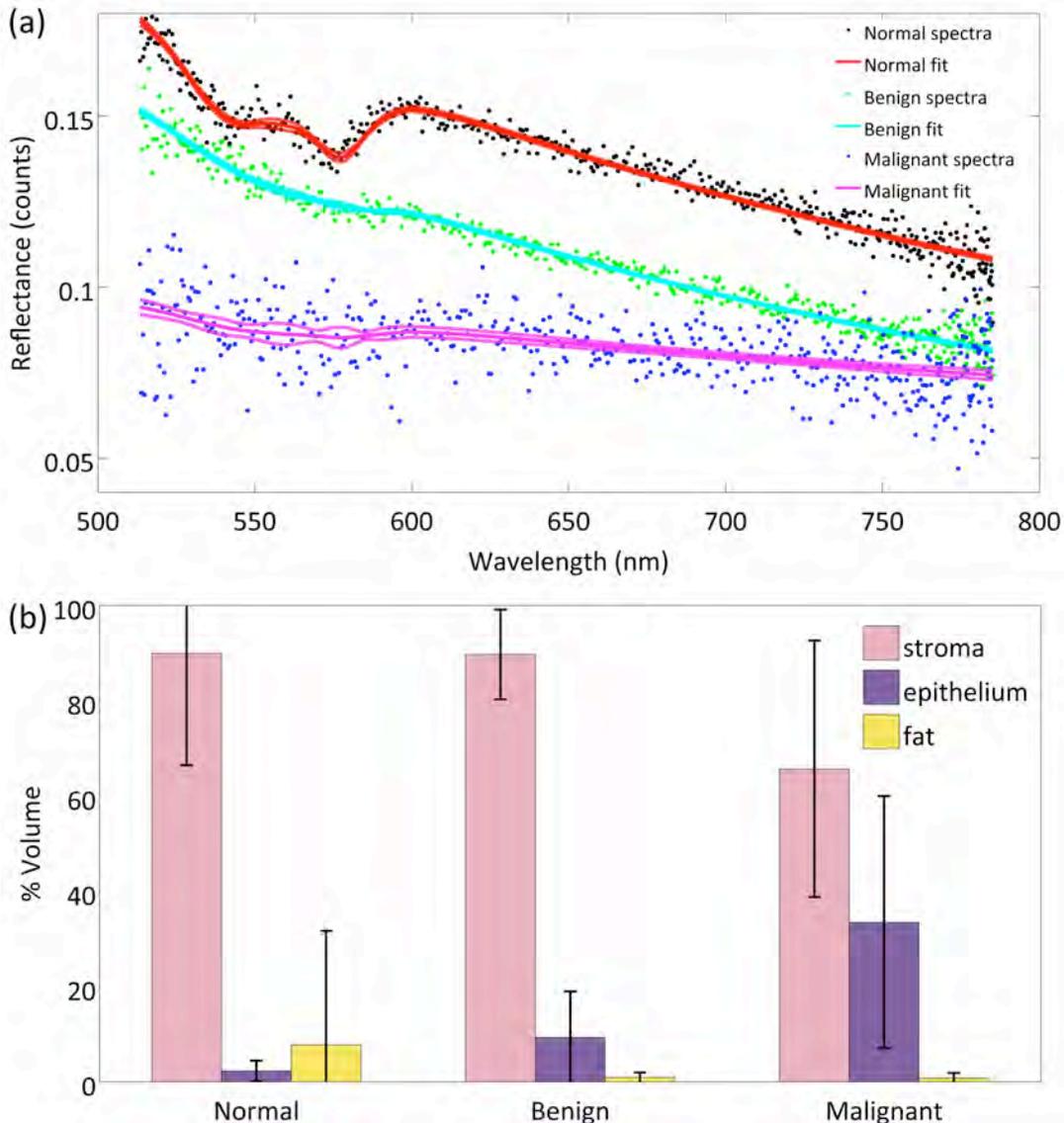
Figure 1(a) illustrates co-registration between the white light image, histology and images of scattering parameters for a tissue sample. Figure 1(b-c) shows box plots of the scattering power and the logarithm of the wavelength-integrated irradiance with outliers removed (those greater than 2 standard deviations from the mean) for all tissue samples. The scattering amplitude is not displayed because it follows the same trend as scattering power per diagnostic category and a correlation is observed between scattering power and logarithm of the scattering amplitude (Figure 1d).



**Figure 1 (a) Co-registration between the digital photograph, histology and maps of scattering power, amplitude, and total-wavelength integrated intensity for a given tissue; (b-c) box plots of relative scatter power and log of the total-wavelength integrated intensity according to diagnosis (outliers > 2 std not displayed); (d) the 3-dimensional features space assembled with scattering parameters and employed by the k-NN classifier.**

Histopathology reveals that the three macroscopic scattering centers found in breast tissue are epithelium, stroma and adipose. Immunohistochemistry shows that the percent distribution of these components varies with diagnosis and registration of scattering maps with pathology illustrate how spectral response changes as a function of diagnosis. This suggests the percent distribution of stroma, epithelium and adipocytes in the effective illumination volume influences scattering response. Standard immunohistochemistry techniques were used to assess the percent distribution of adipose, stroma and epithelium per sample. Formalin fixed and paraffin embedded tissue sections were cut and mounted on OptiPlus™ Positive Charged Barrier slides (BioGenex, San Ramon CA) to test for Anti-Cytokeratins 8 and 18 (clone 5D3; BioGenex, San

Ramon CA. Whole immunostained slides were digitally scanned and montaged using the Surveyor© Automated Specimen Scanning (Objective Imaging Ltd., Cambridge UK) automated stage control bundled software. The epithelium to stroma ratio was quantified using Image-Pro Plus (Media Cybernetics, Bethesda MD) image analysis software. The epithelial and stromal percentages were defined as the percent of CK5D3 positive or Hematoxylin counterstained tissue, thresholded in pseudo-color in the diagnostic regions of interest (ROI), as compared to the total area of the tissue, respectively. Figure 2(a) shows fitted spectra sampled from normal, benign and malignant tissues respectively. Figure 2(b) illustrates how epithelium, stroma and fat content vary between normal, benign and malignant samples based on this analysis.



**Figure 2** In (a) the fitted spectra sampled from normal, benign and malignant tissues respectively, are shown. In (b) the distribution of stroma, epithelium and adipose are shown across the three diagnostic categories classified by immunohistochemistry, for all tissue samples are shown.

Task 2. Develop an automated classification algorithm to provide real-time, un-biased interpretation of scatter images for improved evaluation of breast surgical margins.

### Materials and Methods

#### *2a. k-Nearest Neighbor Classification*

The distribution of scattering parameters demonstrated subtle discrimination between tissue subtypes, but data was multi-parametric and overall classification was challenging. A k-Nearest Neighbor (k-NN) classifier was employed for ready discrimination between tissue pathologies. The k-NN classifier simultaneously interprets multiple scattering parameters for tissue characterization by assigning an unclassified vector (herein referred to as the query point) to the majority diagnosis of its k-nearest vectors found in the feature space. The feature space for three scatter-related parameters (scattering amplitude, scattering power, and total wavelength-integrated intensity) is depicted in Figure 1(d), as well as a query point with unknown diagnosis. All tissue pixels were defined according to a vector in the 3-dimensional feature space and were assigned to the training set (populated feature space) or to the validation set (query points). Sample distributions between training and validation sets were made both randomly and according to a leave-one-out analysis per patient<sup>24</sup>. All training pixels were associated with a known diagnosis according to the pathologist's demarcation of ROIs. The diagnosis of each query point was also determined by the pathologist, but remained unknown to the classifier in order to evaluate its performance.

Additional feature extraction from the actual data set has been shown to compensate for pixel-to-pixel variations and to improve the overall performance of the classifier<sup>25</sup>. Therefore the first four statistical moments (mean, standard deviation, skewness and kurtosis) of each scattering parameter were estimated in a real 2-dimensional spatial window centered about each pixel of interest. These local statistical parameters were concatenated to the actual scatter parameters, and parametric feature space was expanded from 3-dimensions to 15-dimensions. The behavior of the classifier was then studied as a function of two independent variables: the number of nearest neighbors  $k$  and the size of the spatial window used to compute local statistics.

#### *2b. Validation of the Classifier*

In order to optimize the independent variables associated with the classifier, a threefold cross-validation technique was applied for discrimination between not-malignant, malignant and adipose samples and for discrimination between all pathology subtypes identified in Table 1<sup>27, 28</sup>. All data was randomly divided into three non-overlapping sets, with an equal number of pixels per diagnostic category per set. Two of these sets were employed as a training set (used to populate feature space) and the other was employed as a validation set (query points) to compute the classification error. Error was taken to be the percentage of misclassified pixels in the validation set, where a misclassification means that the diagnosis assigned to a pixel by the automated classifier does not match the diagnosis provided by the pathologist. This procedure was repeated three times, for all possible permutations of training and testing sets and the reported classification error is the average of these three executions. This threefold cross-validation was repeated for a varying number of nearest neighbors,  $k$ , and a varying spatial window size for computation of local statistics.

Additionally, leave-one-out analysis was performed per patient, where ROIs from one tissue sample populate the validation set and all other ROI pixels populate the feature space. In this validation procedure, points are not equally distributed between diagnostic categories in either the training or testing sets. Images of the classification results were generated in H&E false color for each tissue sample, allowing one to evaluate whether the predicted diagnosis outside selected ROIs makes sense in context of the entire sample. A mode filter was applied in a sliding window (5x5 pixels) over the k-NN classified image to eliminate impulsive assignment noise. The error and efficacy of the classifier was summarized for all tissue samples. Pixels corresponding to locations where reflectance spectra could not be reliably measured were tagged as masked pixels, and these were excluded from the training and validation sets during all cross-validation procedures.

## Results

Table 2 summarizes the classification efficacy and classification error observed when performing leave-one-out validation for all tissue samples.

	Classification (Not-malignant, Malignant, Fat)			Classification (All Pathologies)						
<b>Classification Error</b>										
<i>Median</i>	8.75			16.8						
<i>Mean</i>	13.0			25.3						
<i>Standard Deviation</i>	13.7			25.0						
<i>Inter-quartile range</i> <i>[min max]</i>	[0 53.5]			[2.15 95.3]						
<b>Total Efficacy</b>	<b>Not- malignant</b>	<b>Malignant</b>	<b>Fat</b>	<b>Normal</b>	<b>Benign</b>	<b>DCIS</b>	<b>IDC</b>	<b>ILC</b>	<b>Inflam</b>	<b>Fat</b>
Accuracy	0.86	0.86	0.98	0.74	0.85	1.00	0.86	0.99	0.99	0.98
Sensitivity	0.90	0.77	0.87	0.74	0.09	0.00	0.77	0.00	0.00	0.87
Specificity	0.82	0.90	0.99	0.74	0.91	1.00	0.90	1.00	1.00	0.99
Negative Predictive Value	0.87	0.88	0.99	0.77	0.92	1.00	0.89	0.99	0.99	0.99
Positive Predictive Value	0.86	0.81	0.95	0.71	0.08	0.00	0.79	0.00	0.00	0.95

**Table 2 Summary of the classification error and total efficacy of the k-NN classifier when discriminating between not-malignant, malignant and adipose tissue and when discriminating between all pathologies. Reported measures based upon ability to discriminate given pathology from all other diagnostic categories evaluated.**

The median classification error is approximately 17% and 9% when discriminating between all pathologies and not-malignant, malignant, and adipose tissue respectively. This is quite close to our performance estimates. When classifying all pathologies, low sensitivity is observed for those classes under-represented in sample space (benign, DCIS, IDC, ILC, inflammation). In any case, the classifier has clinical application because normal epithelium and stroma, invasive ductal carcinoma and adipose are the most frequently encountered tissues during conservative breast surgery. The classifier's sensitivity to not-malignant, malignant and adipose pathologies is 0.90, 0.77 and 0.87 respectively. Sensitivity is lower in malignant samples because its sample population is characterized by greater heterogeneity. While DCIS, IDC and ILC are all considered malignant, morphologically and biologically they are quite distinct. Specificity of the classifier for not-malignant, malignant and adipose pathologies is 0.82, 0.90 and 0.99 respectively. Specificity is lowest in normal tissues because these are characterized by mixed fibro-glandular and adipose content. Epithelial proliferation in malignant tissues was observed to crowd out adipocytes in this study. In a reflectance geometry, scattering from adipocytes results in a very low (noisy) signal. The negative and positive predictive values for each diagnostic category are also reported. These refer to the number of patients with negative and positive results (respectively) who are correctly diagnosed. For surgical margin applications, the surgeon is most interested in a high negative predictive value, ensuring his/her diagnosis of normal or malignant is an accurate one. The negative predictive values for not-malignant and malignant pathologies are 87% and 88% respectively. Although less essential, high positive predictive values prevent any unnecessary re-excisions during surgery.

The confusion matrices in Table 3 illustrate the distribution of misclassified pixels across diagnostic categories when performing leave-one-out analysis for two levels of diagnostic discrimination. This is important to consider because cost to the patient for misclassifying a normal pixel as benign is less than cost to the patient for misclassifying a malignant pixel as normal.

(a) TRUE DIAGNOSIS			(b) TRUE DIAGNOSIS							
		Benign	Malignant	Normal	Benign	DCIS	IDC	ILC	Inflam	Fat
ESTIMATED DIAGNOSIS	Benign	90.8%	23.1%	73.8%	85.2%	83.0%	20.0%	21.3%	3.8%	1.1%
	Malignant	9.2%	76.9%	16.6%	9.2%	8.8%	1.7%	1.3%	0.0%	0.3%
				0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
				9.3%	5.5%	7.2%	77.4%	73.5%	68.1%	9.3%
				0.0%	0.0%	0.0%	0.2%	0.0%	3.4%	0.4%
				0.0%	0.0%	0.0%	0.1%	1.3%	0.0%	1.6%
				0.3%	1.0%	1.0%	0.5%	2.7%	24.8%	87.3%
TOTAL # DIAGNOSED PIXELS		44,000	23,220	31226	5220	194	22547	479	533	7021

Table 3 (a) A confusion matrix, under leave-one-out validation, that represents trends in misclassification and accurate identification of benign and malignant pathologies. (b) A confusion matrix, under leave-one-out validation, that represents trends in misclassification and accurate identification of all pathologies identified by the pathologist. The matrices list the percentage of pixels classified correctly along the diagonal (yellow), and incorrectly off the diagonal. Clear misclassifications are highlighted in gray.

Figure 3 illustrates classification of 6 representative tissue samples. The first column contains a digital photograph of each tissue sample taken immediately after spectral imaging; this is the surgeon’s perspective. Fibro-glandular tissue is white, adipose is yellow-orange and higher concentrations of hemoglobin are red. Histological sections were co-registered to the scattering and white light images and are displayed in column two. Hematoxylin has a deep blue-purple color and stains nucleic acids, which are primarily located in the cell nuclei. Eosin is pink and stains proteins nonspecifically; mainly the cytoplasm and stroma have varying degrees of pink staining. Fat is not preserved during histological processing, so this becomes empty space on the slide. Column 3 illustrates the ROIs identified by the pathologist and they are colored according to their true diagnostic category, while column 4 contains images of the automated diagnosis provided by the k-NN classifier.

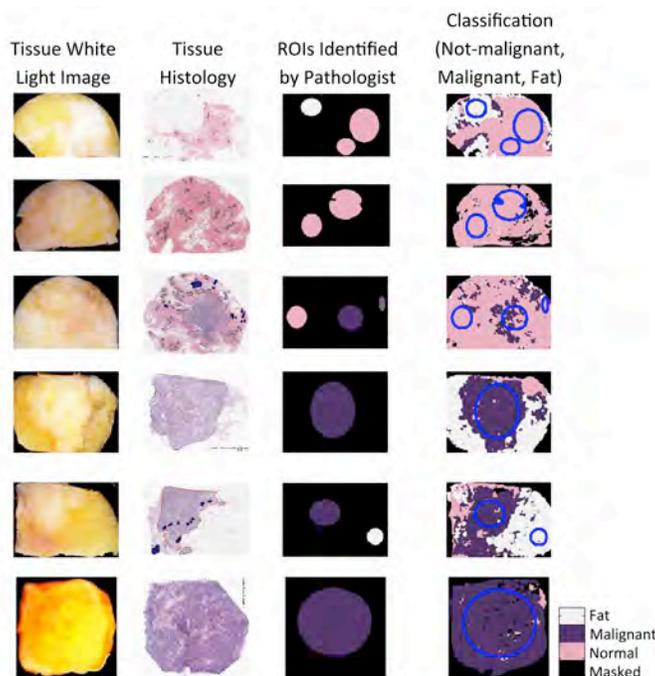


Figure 3 Classification of 6 representative tissue samples. Each row corresponds to a different tissue sample and the following four images are co-registered from left to right: (1) a white light image of the tissue, (2) H&E stained sections of the tissue, (3) true diagnosis of ROIs identified by the pathologist, and (4) classification results when discriminating between not-malignant, malignant and adipose tissues.

We recognize that the surgeon is most interested in a diagnosis of either benign or malignant; therefore the k-NN classifier was executed with this binary level of discrimination. When separating benign and malignant pathologies, the sensitivity and specificity was 91% and 77% respectively; and the system achieved positive and negative predictive values of 88% and 81% respectively.

### Discussion

This work demonstrates that morphological features pertinent to a tissue's diagnosis may be ascertained from confocal detection of broadband reflectance, with a mesoscopic resolution that permits scanning of an entire margin for residual disease. The technical aspects and optimization of a k-NN classifier for automated diagnosis of pathologies is presented and validated in 29 specimens of breast tissue. The classifier's discriminating capabilities improved with the inclusion of local statistics, likely accounting for microscopic tissue heterogeneities. Initially, discrimination between all pathologies identified by WAW was attempted; however, inadequate sampling of uncommon pathologies rendered their classification less robust. Given the sample population, discrimination between not-malignant, malignant and adipose pathologies was most intuitive; particularly because these diagnostic categories correspond to the three macroscopic scattering centers in breast tissue (stroma, epithelium and adipocytes). Negative predict values of 87%, 88% and 99% were achieved for not-malignant, malignant and adipose tissues respectively. In the same order, their positive predictive values were 86%, 81% and 95%. The classifier was most sensitive to not-malignant and adipose tissues because the malignant population was pathologically very diverse; including samples of invasive lobular carcinoma, invasive ductal carcinoma and ductal carcinoma in situ. Specificity was lowest in not-malignant samples because of their mixed fibro-glandular and adipose content. In the context of conservative surgery, the goal of treatment is to maximize removal of malignant tissues while minimizing damage to healthy, viable tissue. When discriminating between benign and malignant tissues only, a sensitivity of 91% and a specificity of 77% was achieved. This sensitivity is significantly higher than those reported for frozen section analysis (which is not practical for lumpectomy margins because of the problems associated with freezing and cutting adipose tissues) and diffuse reflectance spectroscopy; although its specificity is lower<sup>29-31</sup>. Even though overall efficacy of the classifier exceeds or is comparable to other intra-operative assessment techniques, integration of spectroscopy methods into the surgical suite will require a better negative predictive value, ensuring that the surgeon's diagnosis of benign or malignant is an accurate one. Additionally, higher positive predictive values would prevent unnecessary tissue removal during surgery.

Improvement of the classifier's performance may fundamentally be achieved with greater sampling. As the number of data points in feature space increases, so does the accuracy of the classifier [Fukunaga, 1972 #2008]. Particularly, the classifier's sensitivity to malignant pathologies has the most to gain and could be improved with equal representation of IDC, ILC and DCIS in feature space. As feature space expands, so will computational costs. Rather than calculating the distance between each query point and every point in feature space, a KD-tree may be employed to optimize the search algorithm<sup>32</sup>. The classifier was trained with ROIs obviously belonging to a diagnosis - normal and malignant pathologies were identified by cellular features in fibro-glandular regions and adipose ROIs were measured far from any fibro-glandular tissue. Perhaps the performance of the classifier would improve if an additional level of classification was employed, so that each diagnostic category was also labeled according to the subtypes, 'fibro-glandular' or 'fatty-fibro-glandular'.

While mechanically scanning the sample was time intensive and therefore not suitable for clinical translation, a second-generation system has been developed that employs scanning-beam architecture to image tissue fields up to 1x1cm<sup>2</sup> within 9-12 minutes. The high-throughput platform combines a broadband telecentric scanning design with dark-field illumination/detection optical path to allow efficient rejection of specular light while maintaining a consistent sampling geometry across the entire imaging field. System details can be found in the reference<sup>33</sup>; and this system is now used to image breast lumpectomy specimens. To expand upon the variety of data collected, a supercontinuum white light source is used with the new scanning-beam spectral imager. This source allows for broadband spectral imaging of breast pathology in a waveband

(400-750nm) that is sensitive to both tissue morphology and biochemical composition. Particularly, within this waveband one may determine the concentration of oxygenated and deoxygenated hemoglobin, beta carotene and blood break-down products in a tissue, while simultaneously extracting scattering features in a region with minimal absorption (beyond 620nm). Addition of other optical parameters to the classifier is extremely simple, only involving an expansion of feature space (update the vector describing each pixel to N-dimensions, where N is the number of parameters). It would be particularly useful to generate a comprehensive dataset with parameters describing all possible endogenous light-tissue interactions (scattering, absorption, fluorescence), so that the most diagnostically discriminating and robust parameters could be identified and optimized during data collection. Note that beta-carotene is a member of the carotenoids and gives fat its highly pigmented, yellow color; we hope that its absorption spectra will improve classification of tissues with high adipose content.

Finally, understanding the relationship between the optical and biological properties of a tissue will ultimately improve the diagnostic utility of optical techniques – permitting optimization of the measurement procedure and signal analysis for enhanced sensitivity to differentiating features. The technique remains to be tested intra-operatively; future clinical studies will reveal how the system may enhance existing surgery and pathologic procedures.

**Key Research Accomplishments**

- Validation and optimization of a k-Nearest Neighbor classifier to automatically detect breast tissue pathologies based upon direct sampling of the scattering spectral response.
- Assessment of classifier performance in 67,000 spectra from 29 breast tissue specimens. When discriminating between benign and malignant pathologies, a sensitivity and specificity of 91% and 77% were achieved respectively.
- Detailed sub-tissue analysis was performed to consider how diverse pathologies influence scattering response and overall classification efficacy.
- Development of a dark-field spectral imaging system for rapid scanning of thick tissue specimens over a 1cm<sup>2</sup> field of view.

## Reportable Outcomes

### Manuscripts

#### *Peer-reviewed publications*

- A. M. Laughney**, Krishnaswamy, V., Garcia-Allende, P.B., Conde, O.M., Wells, W.A., Paulsen, K.D., Pogue, B.W., "Automated classification of breast pathology using local measures of broadband reflectance," *JOURNAL OF BIOMEDICAL OPTICS* (2010).
- V. Krishnaswamy, **A. M. Laughney**, K. Paulsen, and B. Pogue, "A Dark-Field In Situ Scanning Spectroscopy Platform for Broadband Imaging of Resected Tissue," *Optics Letters* (accepted).

#### *Conference Publications*

- Laughney, A.**, Krishnaswamy, V., Wells, W., Paulsen, K., Pogue, B., "Optical assessment of pathology in surgically resected breast cancers" *Proceedings of SPIE BiOS Conference, Biomedical Applications of Light Scattering V, San Francisco CA, January 2011.*
- Laughney, A.**, Krishnaswamy, V., Garcia-Allende, P.B., Wells, W.A., Conde, O.M., Paulsen, K.D., Pogue, B.W., "Imaging Breast Pathology in situ using Broadband Scatter Spectroscopy and a K-Nearest Neighbor Classifier," *Proceedings of the OSA BIOMED Conference, Bio-optics in clinical application, Miami FL, April 2010.*

### Presentations

"A spatially-modulated scatter imaging system to detect tumor-associated stroma," *Breast Tumor Board Conference at Dartmouth-Hitchcock Medical Center, Spring 2011.*

"Optical characterization of pathologies in surgically resected breast cancers," *Biomedical Lecture Series, Thayer, Spring 2011.*

"System design for optical characterization of tissue pathologies in situ," *NSF ADVANCE Workshop at Rice University, Fall 2010.*

"Optimizing spectral contrast in breast lesions relative to normal tissue for surgical guidance," *Thayer, Spring 2010.*

"Automated Classification of Breast Pathology using Local Measures of Broadband Reflectance," *Optical Society of America 2010 BIOMED Conference (Miami, FL), Spring 2010; Beckman Institute at the University of California Irvine, Winter 2009.*

### Research Opportunities Received

2011 Biophotonics International Graduate Summer School  
NSF ADVANCE Workshop at Rice University  
Pathobiology of Cancer Workshop by the American Association of Cancer Research  
Neukom Institute Travel Grant

## Conclusions

In this contribution, we validate and optimize the ability of a k-NN classifier to automatically detect breast tissue pathologies based upon sampling of scattering spectral features. The sampling volume was specifically chosen to be sensitive to architectural changes addressed by a pathologist during microscopic assessment of a surgical specimen, while also permitting its wide field scanning. Performance of the classifier was assessed in 29 breast tissue specimens, and when discriminating between benign and malignant pathologies, a sensitivity and specificity of 91% and 77% was achieved. Further, detailed sub-tissue analysis was performed to consider how diverse pathologies influence scattering response and overall scattering efficacy. The purpose of automating classification of scattering response from diagnostic pathologies is to assess involved surgical margins for cancer during primary surgery. Identification of residual disease at the time of primary surgery offers clear value to the patient by decreasing re-excision rates and improving a patient's survival advantage. If residual tumor is present at one or more margins, the surgeon, before closing, could be advised to remove more tissue immediately, rather than a later re-excision. Optical characterization of tissue is expected to improve completeness of resection during breast conserving surgery because molecular interaction with light provides specific information about a tissue's biochemistry and organelle morphology, which are altered by disease.

## References

1. U. Veronesi, N. Cascinelli, L. Mariani, M. Greco, R. Saccozzi, A. Luini, M. Aguilar and E. Marubini, "Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer," *New Engl J Med* 347 (16), 1227-1232 (2002).
2. R. Arriagada, M. G. Le, F. Rochard and G. Contesso, "Conservative treatment versus mastectomy in early breast cancer: Patterns of failure with 15 years of follow-up data," *J Clin Oncol* 14 (5), 1558-1564 (1996).
3. J. A. Jacobson, et al., "Ten-Year Results of a Comparison of Conservation with Mastectomy in the Treatment of Stage I and II Breast Cancer," *New Engl J Med* 332 (14), 907-911 (1995).
4. J. A. van Dongen, et al., "Long-term results of a randomized trial comparing breast-conserving therapy with mastectomy: European Organization for Research and Treatment of Cancer 10801 trial," *J Natl Cancer I* 92 (14), 1143-1150 (2000).
5. M. Blichert-Toft, C. Rose, J. A. Andersen, M. Overgaard, C. K. Axelsson, K. W. Andersen and H. T. Mouridsen, "Danish randomized trial comparing breast conservation therapy with mastectomy: six years of life-table analysis. Danish Breast Cancer Cooperative Group," *J Natl Cancer Inst Monogr* (11), 19-25 (1992).
6. J. Q. Brown, et al., "Optical Assessment of Tumor Resection Margins in the Breast," *IEEE JOURNAL OF SELECTED TOPICS IN QUANTUM ELECTRONICS* 16 (3), 530-544 (2010).
7. S. A. Khan and F. Eladoumikdachi, "Optimal Surgical Treatment of Breast Cancer: Implications for Local Control and Survival," *Journal of Surgical Oncology* 101 (8), 677-686 (2010).
8. M. Kaufmann, M. Morrow, G. von Minckwitz, J. R. Harris and M. Biedenkopf Expert Panel, "Locoregional Treatment of Primary Breast Cancer Consensus Recommendations From an International Expert Panel," *Cancer* 116 (5), 1184-1191 (2010).
9. S. Kennedy, et al., "Optical breast cancer margin assessment: an observational study of the effects of tissue heterogeneity on optical contrast," *BREAST CANCER RESEARCH* 12 (6), R91 (2010).
10. T. Bydlon, et al., "Performance metrics of an optical spectral imaging system for intra-operative assessment of breast tumor margins," *Opt Express* 18, 8058 - 8076 (2010).
11. T. M. Bydlon, et al., "Rapid Optical Imaging of Breast Tumor Margins: Final Results from a 100-Patient Clinical Study," *Cancer Res* 69 (24), 770s-771s (2009).
12. V. Krishnaswamy, Wells, W.A., Laughney, A.M., Samkoe, K.S., Pogue, B.W., presented at the SPIE Biomedical Applications of Light Scattering III, San Jose, CA, USA, 2009 (unpublished).
13. A. M. Laughney, V. Krishnaswamy, P. B. Garcia-Allende, O. M. Conde, W. A. Wells, K. D. Paulsen and B. W. Pogue, "Automated classification of breast pathology using local measures of broadband reflectance," *J Biomed Opt* 15 (6), 066019 (2010).
14. A. Nabavi, H. Thurm, B. Zountsas, T. Pietsch, H. Lanfermann, U. Pichlmeier, M. Mehdorn and -. A. R. Glioma, "Five-Aminolevulinic Acid for Fluorescence-Guided Resection of Recurrent Malignant Gliomas: A Phase II Study," *Neurosurgery* 65 (6), 1070-1076 (2009).
15. V. Krishnaswamy, P. J. Hoopes, K. S. Samkoe, J. A. O'Hara, T. Hasan and B. W. Pogue, "Quantitative imaging of scattering changes associated with epithelial proliferation, necrosis, and fibrosis in tumors using microsampling reflectance spectroscopy," *JOURNAL OF BIOMEDICAL OPTICS* 14 (1), - (2009).
16. B. W. Pogue and T. Hasan, "Fluorophore quantitation in tissue-simulating media with confocal detection," *IEEE JOURNAL OF SELECTED TOPICS IN QUANTUM ELECTRONICS* 2 (4), 959-964 (1996).
17. B. W. Pogue and G. Burke, "Fiber-optic bundle design for quantitative fluorescence measurement from tissue," *APPLIED OPTICS* 37 (31), 7429-7436 (1998).
18. J. R. Mourant, J. P. Freyer, A. H. Hielscher, A. A. Eick, D. Shen and T. M. Johnson, "Mechanisms of light scattering from biological cells relevant to noninvasive optical-tissue diagnostics," *APPLIED OPTICS* 37 (16), 3586-3593 (1998).

19. H. J. Vanstaveren, C. J. M. Moes, J. Vanmarle, S. A. Prahl and M. J. C. Vangemert, "Light-Scattering in Intralipid-10-Percent in the Wavelength Range of 400-1100 Nm," *APPLIED OPTICS* 30 (31), 4507-4514 (1991).
20. M. Bartek, X. Wang, W. Wells, K. D. Paulsen and B. W. Pogue, "Estimation of subcellular particle size histograms with electron microscopy for prediction of optical scattering in breast tissue," *JOURNAL OF BIOMEDICAL OPTICS* 11 (6), 064007 (2006).
21. X. Wang, et al., "Image reconstruction of effective Mie scattering parameters of breast tissue in vivo with near-infrared tomography," *JOURNAL OF BIOMEDICAL OPTICS* 11 (4), - (2006).
22. X. Wang, B. W. Pogue, S. D. Jiang, X. M. Song, K. D. Paulsen, C. Kogel, S. P. Poplack and W. A. Wells, "Approximation of Mie scattering parameters in near-infrared tomography of normal breast tissue in vivo," *JOURNAL OF BIOMEDICAL OPTICS* 10 (5), 051704 (2005).
23. S. L. Jacques, Prahl, S., (Oregon Medical Laser Center, Portland, OR, 2010).
24. O. Eytan, B. A. Sela and A. Katzir, "Fiber-optic evanescent-wave spectroscopy and neural networks: application to chemical blood analysis," *APPLIED OPTICS* 39 (19), 3357-3360 (2000).
25. P. B. Garcia-Allende, V. Krishnaswamy, P. J. Hoopes, K. S. Samkoe, O. M. Conde and B. W. Pogue, "Automated identification of tumor microscopic morphology based on macroscopically measured scatter signatures," *JOURNAL OF BIOMEDICAL OPTICS* 14 (3), - (2009).
26. D. A. Burns, Ciurczak, E.W., *Handbook of Near-Infrared Analysis* (Chpt. 15), 3rd ed. (CRC Press, 2008).
27. C. Goutte, "Note on free lunches and cross-validation," *Neural Comput* 9 (6), 1245-1249 (1997).
28. H. Y. Zhu and R. Rohwer, "No free lunch for cross-validation," *Neural Comput* 8 (7), 1421-1426 (1996).
29. N. Cabioglu, et al., "Role for Intraoperative Margin Assessment in Patients Undergoing Breast-Conserving Surgery," *Annals of Surgical Oncology* 14 (4), 1458-1471 (2007).
30. T. M. Breslin, G. Palmer, C. Fang, K. W. Gilchrist, F. Xu and N. Ramanujam, "Auto fluorescence and diffuse reflectance properties of malignant and benign breast tissues," *Annals of Surgical Oncology* 10 (1), S16-S16 (2003).
31. G. M. Palmer, C. F. Zhu, T. M. Breslin, F. S. Xu, K. W. Gilchrist and N. Ramanujam, "Comparison of multiexcitation fluorescence and diffuse reflectance spectroscopy for the diagnosis of breast cancer (March 2003)," *IEEE T Bio-Med Eng* 50 (11), 1233-1242 (2003).
32. K. Fukunaga, *Introduction to Statistical Pattern Recognition*. (Academic Press, New York, 1972).
33. V. Krishnaswamy, Laughney, A.M., Paulsen, K.D., Pogue, B.W., "A Dark-Field In Situ Scanning Spectroscopy Platform for Broadband Imaging of Resected Tissue," *OPTICS LETTERS* (2010 (in review)).
34. A. Mahadevan-Jansen, M. F. Mitchell, N. Ramanujam, A. Malpica, S. Thomsen, U. Utzinger and R. Richards-Kortum, "Near-infrared Raman spectroscopy for in vitro detection of cervical precancers," *Photochem Photobiol* 68 (1), 123-132 (1998).
35. N. Q. Mirza, et al., "Ductal carcinoma-in-situ: Long-term results of breast-conserving therapy," *Annals of Surgical Oncology* 7 (9), 656-664 (2000).
36. E. B. Claus, S. Petruzella, E. Matloff and D. Carter, "Prevalence of BRCA1 and BRCA2 mutations in women diagnosed with ductal carcinoma in situ," *Jama-Journal of the American Medical Association* 293 (8), 964-969 (2005).
37. N. S. Goldstein, L. Kestin and F. Vicini, "Intraductal carcinoma of the breast - Pathologic features associated with local recurrence in patients treated with breast-conserving therapy," *American Journal of Surgical Pathology* 24 (8), 1058-1067 (2000).
38. M. R. Bani, M. P. Lux, K. Heusinger, E. Wenkel, A. Magener, R. Schulz-Wendtland, M. W. Beckmann and P. A. Fasching, "Factors correlating with reexcision after breast-conserving therapy," *Ejso* 35 (1), 32-37 (2009).

39. P. B. Garcia-Allende, V. Krishnaswamy, P. J. Hoopes, K. S. Samkoe, O. M. Conde and B. W. Pogue, "Automated identification of tumor microscopic morphology based on macroscopically measured scatter signatures," JOURNAL OF BIOMEDICAL OPTICS 14 (3) (2009).