Development of a spatially offset Raman spectroscopy probe for breast tumor surgical margin evaluation

Matthew D. Keller
Elizabeth Vargis
Nara de Matos Granja
Robert H. Wilson
Mary-Ann Mycek
Mark C. Kelley
Anita Mahadevan-Jansen

Development of a spatially offset Raman spectroscopy probe for breast tumor surgical margin evaluation

Matthew D. Keller,[a] Elizabeth Vargis,[a] Nara de Matos Granja,[b] Robert H. Wilson,[c] Mary-Ann Mycek,[c,d,e] Mark C. Kelley,[f] and Anita Mahadevan-Jansen[a]
[a]Vanderbilt University, Department of Biomedical Engineering, Nashville, Tennessee 37235
[b]Vanderbilt University Medical Center, Division of Pathology, Nashville, Tennessee 37232
[c]University of Michigan, Applied Physics Program, Ann Arbor, Michigan 48109
[d]University of Michigan, Department of Biomedical Engineering, Ann Arbor, Michigan 48109
[e]University of Michigan, Comprehensive Cancer Center, Ann Arbor, Michigan 48109
[f]Vanderbilt University Medical Center, Division of Surgical Oncology, Nashville, Tennessee 37232

Abstract. The risk of local recurrence for breast cancers is strongly correlated with the presence of a tumor within 1 to 2 mm of the surgical margin on the excised specimen. Previous experimental and theoretical results suggest that spatially offset Raman spectroscopy (SORS) holds much promise for intraoperative margin analysis. Based on simulation predictions for signal-to-noise ratio differences among varying spatial offsets, a SORS probe with multiple source-detector offsets was designed and tested. It was then employed to acquire spectra from 35 frozen–thawed breast tissue samples *in vitro*. Spectra from each detector ring were averaged to create a composite spectrum with biochemical information covering the entire range from the tissue surface to ∼2 mm below the surface, and a probabilistic classification scheme was used to classify these composite spectra as "negative" or "positive" margins. This discrimination was performed with 95% sensitivity and 100% specificity, or with 100% positive predictive value and 94% negative predictive value. © *2011 Society of Photo-Optical Instrumentation Engineers (SPIE).* [DOI: 10.1117/1.3600708]

Keywords: spatially offset Raman spectroscopy; breast cancer; margin evaluation; Monte Carlo simulation.

1 Introduction

For many of the approximately 180,000 women diagnosed with early stage invasive breast cancer or carcinoma *in situ* each year,[1] a viable treatment option is breast conserving therapy (BCT). The surgical portion of BCT involves a partial mastectomy, or lumpectomy, to remove only the primary lesion with a small amount of surrounding normal tissue.[2] Depending on the hospital, the depth of normal tissue required from the surgical margin (i.e., the surface of the excised specimen) to the tumor is typically 1 to 2 millimeters.[3] This situation is illustrated in Fig. 1; if a sufficient amount of normal tissue exists, as in the right side of Fig. 1, the margins are said to be negative for tumor. If tumor-positive margins are found, as shown on the left side of Fig. 1, a second operation is necessary because positive margins are a major predictor of local tumor recurrence.[4] Currently available intraoperative margin evaluation tools, such as simple visual examination, ultrasound, cytological examination ("touch prep"), and frozen section analysis, all have significant drawbacks in terms of accuracy and/or time required,[5–8] so there is a need for an automated, real-time method to accurately evaluate surgical margins during BCT.

Several optical techniques have recently been investigated for breast tumor margin evaluation. A diffuse reflectance imaging system[9–11] was employed to assess 55 tumor margins from 48 patients, and classified whole margins as negative versus positive/close with 79% sensitivity and 67% specificity.[11] Kennedy et al. employed the same diffuse reflectance spectroscopy system to characterize tumor margins from 100 patients.[12] Keller et al. used combined diffuse reflectance and autofluorescence spectroscopy to classify individual points from 32 patients on margins as negative versus positive with 85% sensitivity and 96% specificity, and also demonstrated the ability to perform autofluorescence spectral imaging of larger regions.[13] Nguyen et al. used optical coherence tomography (OCT) images from 20 patients to classify the margin status with 100% sensitivity and 82% specificity.[14]

In the one report of using Raman spectroscopy for a margin analysis tool,[15] measurements were made *in vivo* rather than on the excised specimen; the latter approach is the current standard practice in surgical pathology. That study also used a conventional fiber optic probe with source and detector fibers adjacent to each other, which allows only limited depth sensing. More recently, spatially offset Raman spectroscopy (SORS) has been shown to be a reliable method for recovering biological Raman spectra from depths greater than those possible with standard techniques.[16–24] It does so because detection elements spaced radially further from source elements are more sensitive to photons traveling deeper beneath the tissue surface and to greater radial distances due to multiple scattering (demonstrated in Fig. 1). We have previously demonstrated the ability of SORS to detect spectral contributions from breast tumors buried under 0.5 to 2 mm of normal breast tissue.[16] In addition, a SORS Monte Carlo (MC) code was developed to quantify signals obtained

Address all correspondence to: Anita Mahadevan-Jansen, Vanderbilt University, Department of Biomedical Engineering, Nashville, Tennessee 37235; Tel: 605-343-4787; Fax: 615-343-7919; E-mail: anita.mahadevan-jansen@vanderbilt.edu.
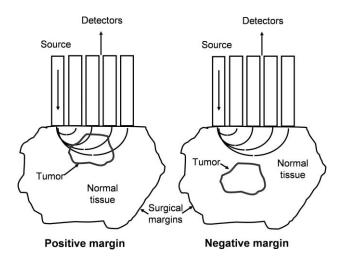
Fig. 1 Drawing of tumor-positive versus tumor-negative margins, defined by the distance between the surgical margin and the tumor boundary. Overlaid are general photon migration paths demonstrating the advantage of SORS for this application.

Fig. 2 Simulation results for total number of Raman photons detected as a function of detector S-D offset, normalized to a maximum of 1. Results were averaged for a variety of thicknesses of the top two tissue layers (0.5 to 2 mm for top normal layer, 0.1 to 20 mm for middle tumor layer). No error bars are shown because the standard deviation was less than 1%.

from layered constructs of normal breast tissue overlying breast tumors.[25] In particular, the code was used to examine the effects of layer thicknesses and overall geometries on relative tumor contributions to detected spectra for a range of source-detector (S-D) offsets.[25] To detect a tumor signature within the first 2 mm from the surface, the resulting spectrum at a given S-D offset must contain at least a 5% contribution from the tumor. To achieve this level of contribution, it was found that the tumor would have to be ~0.1-mm thick under 0.5 mm of normal tissue, or ~1-mm thick under 2-mm of normal tissue.[25]

Combining the results of experimental[16] and numerically simulated[25] SORS indicated that to be sensitive to breast tumors located up to 2 mm beneath normal breast tissue, as needed for surgical margin evaluation, a maximum S-D offset of ~3.5 mm should be used. With larger offsets, the measurements could possibly detect large tumors from over 2 mm deep and create false positives; also, recording spectra with adequate signal-to-noise ratios (SNRs) at larger offsets is difficult since fewer photons tend to escape the tissue surface as the S-D offset increases. In addition, a shortcoming of the previous experimental work was the need to translate the single detector fiber for each measurement. Thus, the goal of this work was to design, test, and implement a multiseparation SORS probe for breast tumor surgical margin evaluation. In particular, this manuscript describes the use of the previously developed SORS Monte Carlo code to investigate the theoretical drop in SNR as a function of the S-D offset, the design of a SORS probe based on the above theoretical and experimental findings, its testing to ensure comparable signal quality in each detector ring, and its use in acquiring spectra from breast cancer samples to assess its ability to accurately evaluate surgical margin status.

2 Materials and Methods

2.1 *SORS Probe Design*

The primary criterion for designing a SORS probe for breast tumor margin analysis was to ensure proper depth sampling—that is, to develop a probe sens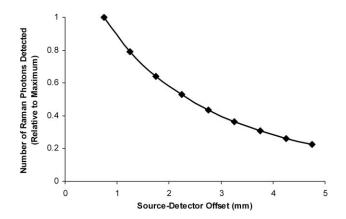itive to tumor spectral signatures if the tumor is anywhere within the first 2 mm in depth from the excised surface. As noted, the relevant S-D offsets for this purpose were determined to be < 3.5 mm. To investigate the drop in SNR as S-D offset increases, SORS Monte Carlo simulations were run using the same model as for the previous results.[25] The Raman MC code is a modified version of a multilayer, multifluorophore code developed by Vishwanath and Mycek.[26–28] As previously described,[25] the code uses a typical Monte Carlo implementation to simulate photon propagation and detection. It also defines a probability that a given excitation photon will be Raman scattered, according to the tissue layer's properties, at each propagation step. Raman scattered photons are then propagated according to new optical properties of each tissue layer, and those Raman photons exiting the top surface of the tissue construct within the collection cone of a given detector bin are recorded. Simulations were run for 3-layered samples, consisting of a top layer of 0.5, 1, or 2 mm of normal breast tissue, a 0.1 to 20-mm thick middle layer of breast tumor, and then a 2-cm thick bottom layer of normal breast tissue to mimic the clinical situation of semi-infinite geometry.

As a metric for SNR, the total number of simulated Raman photons, originating from any layer, reaching each detector bin was counted and normalized to a maximum of 1, since we are only interested in how SNR falls off with the S-D offset. Because the raw signal strength was consistent among the 4 rings, it was assumed that noise levels were consistent among the 4 rings as well. Figure 2 shows the mean of these SNR curves; since the standard deviation was less than 1% over the range of thicknesses for the top two layers, no error bars are shown. As predicted, the number of Raman photons detected fell off at what appears to be an exponential decay as a function of the S-D offset. Although the trend of SNR as a function of S-D offset was consistent regardless of relative thicknesses of normal and cancerous breast tissues, the trend may not hold for other tissue types, especially those in which optical properties can vary more drastically in inhomogeneous regions.

These results were used to aid in the design of a multiseparation SORS probe (assembled by EMVision, Loxahatchee, Florida), whose distal tip is shown in Fig. 3. A single 400-μm
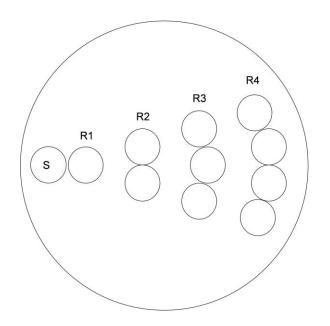
Fig. 3 Schematic of tip of SORS probe. S—source fiber, all other circles—collection fibers. Each ring of collection fibers is separated by 1 mm (center to center) from the previous ring, with 0.5 mm between the centers of the source fiber and the first detector ring (R1).

diameter source, or excitation fiber, is found on one end and 4 (partial) rings of 300-μm diameter collection fibers extend radially outward. The excitation fiber includes a bandpass filter at its tip to narrow the laser line, and the collection fibers have long-pass filters at their tips to reject elastically scattered light. The center to center distances of the excitation fiber to each detection ring are 0.5, 1.5, 2.5, and 3.5 mm. Based on Fig. 2, an additional collection fiber was added to each consecutive detector ring to make the SNRs from each ring more comparable to one another. While the curve in Fig. 2 is not linear, adding a single fiber for each larger-offset ring provided the closest approximation of equivalent SNRs, if all of the fibers in each ring were binned. Adding a fiber in each successive detector ring had an added benefit of increasing the lateral sampling volume of the probe as well.

2.2 Instrumentation and Data Processing

The SORS probe delivered 80 mW of power from a 785 nm diode laser (I0785MB0350M, Innovative Photonics Solutions, Monmouth Junction, New Jersey). The collection fibers delivered light to a near-infrared-optimized spectrograph (LS785, Princeton Instruments, Princeton, New Jersey), which dispersed the light to be recorded by a deep depletion, thermo-electrically cooled CCD (Pixis 400BR, Princeton Instruments).

Each acquisition with the SORS probe recorded 4 spectra—one from each detector ring. Each ring was calibrated separately since the inherent curvature in the detection system created slight but noticeable differences in peak locations on the CCD among different rings. A neon-argon lamp, naphthalene, and acetaminophen standards were used to calibrate the wavenumber axis, and a National Instrument Standards and Tecnology-calibrated tungsten-halogen lamp was used to correct the system response.[29] After wavenumber binning (in 3.5 cm^{-1} steps given the system resolution of $\sim$7 cm^{-1}) and noise smoothing,

the background fluorescence was subtracted with a modified polynomial fit algorithm,[29] and the spectra were normalized according to their overall mean intensities. To create a composite spectrum with equal weighting from all 4 rings, which would contain information from the entire 2 mm sampling depth, the binned spectra from each of the 4 rings were averaged after processing.

2.3 SNR Testing

To ensure the probe's ability to gather spectra from each ring with comparable SNRs, spectra were acquired for 20 s each from 12 different spots on a $\sim$1 cm thick piece of chicken breast (muscle). The spectra were processed as described in Sec. 2.2, and the SNR of the binned spectrum from each ring was calculated by dividing the height of the 1445 cm^{-1} peak, which is the strongest peak in all samples measured, by the standard deviation of the flat (i.e., no Raman signal) spectral range between the 1656 and 1750 cm^{-1} peaks, which represent the noise inherent in the system that could not be removed via pre- or post-processing.

2.4 In vitro Sample Measurements

With approval by the Vanderbilt Institutional Review Board (No. 050551) and the U.S. Army Medical Research and Materiel Command's Human Research Protection Office, fresh frozen human breast tissue samples were acquired from the Cooperative Human Tissue Network. Frozen-thawed tissues are not perfect surrogates for freshly excised tissue, as their optical properties can differ from each other.[30] A recent study by Reble et al.[31] demonstrated that Raman sampling volumes can vary substantially based on a tissue's optical properties, especially the reduced scattering coefficient. Nevertheless, using such tissues is a common first step for breast cancer studies,[30,32] and a recent study showed nearly an equivalent performance of an algorithm for differentiating normal, benign, and malignant breast tissues developed with *in vitro* Raman spectra and applied to *in vivo* measurements.[33]

In total, 35 samples were included in the study; 15 samples had either no tumor ($n = 13$) or a tumor > 2 mm beneath normal tissue at the point of measurement ($n = 2$), and were thus labeled as "negative margins," while 20 samples had a tumor [15 invasive ductal carcinoma (IDC) and 5 invasive lobular carcinoma (ILC)] regions within the first 2 mm from the measurement surface, and were thus labeled "positive margins." Of the negative samples, 7 were predominantly adipose and 8 were varying compositions of adipose and fibroglandular tissue. Of the positive samples, 8 had tumors underlying various compositions of normal tissue ranging from 0.1 to 1.5 mm thick, and 12 samples had tumor regions at the surface under at least part of the probe. Wherever possible, measurements from the tumor samples were taken such that the SORS probe was placed on a small region of visually normal-appearing tissue on top of the actual tumor to mimic the situation of margin evaluation. Spectra were recorded for 10 to 30 s and processed as described above. Measurement sites were inked, fixed in formalin, and serially sectioned to correlate the spectra with histopathology diagnoses of tissue type and the depths from the measurement surface of those tissues.

In this manner, the analysis was done to discriminate "negative" margins from "positive" margins.

2.5 *Classification of Margin Status*

The composite spectrum from averaging all 4 detector rings was used for analysis, and if there were histological evidence of tumor cells within 2 mm of the measurement surface, the "margin" was considered positive. All tumor-positive measurements were lumped into a single category since the surgeon simply needs to know whether any types of malignant cells remain too close to the margin. Discrimination was performed with sparse multinomial logistic regression (SMLR),[34] a Bayesian machine-learning framework that computes the posterior probability of a spectrum belonging to each tissue class based on a labeled training set. In the case of this binary analysis, whichever class had the higher probability of membership was the one to which the spectrum was classified. SMLR also includes inherent dimensionality reduction as it seeks to create sparse basis vectors, which is important for these data sets given their small sizes. Since each *in vitro* sample had only one measurement site (their sizes relative to the probe precluded multiple independent sites), SMLR was run with leave-one-out cross-validation. A range of input parameters to SMLR were tested, and the combination that provided the most accurate classification, while also maximizing sparsity, was using a Laplacian prior, direct kernel, lambda value of 0.01, and not adding a bias term.

3 Results

Figure 4 shows the results of the SNR testing on the chicken muscle. Rings 1 and 4 of the SORS probe, with 1 and 4 fibers per ring, and with S-D offsets of 0.5 and 3.5 mm, respectively, displayed nearly identical SNRs. Rings 2 and 3 showed smaller SNRs compared with ring 1, but only by ~30% and 20%, respectively. This trend was expected based on the shape of Fig. 2, although the signal strengths of rings 2 and 3 were smaller than predicted by the simulations. The likely reason is that when imaging the detection fibers for the two middle rings during alignment and testing, their throughput appeared to be lower, compared with the fibers for rings 1 and 4. Even so, the design
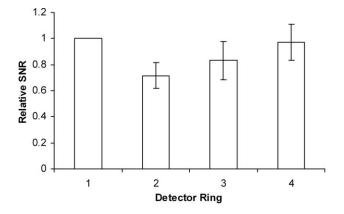
Fig. 5 Typical normalized composite spectra from SORS probe of normal breast tissue versus malignant breast tumor tissue.

of the SORS probe effectively accounted for SNR fall-off with increasing S-D offset.

Figure 5 shows typical composite spectra recorded from pure normal breast tissue and a pure breast tumor (invasive ductal carcinoma for this example; other tumors have similar spectra[35]) tissue with the SORS probe. As in the previous study,[16] there are numerous spectral regions with major differences between the two tissue types. In particular, tumor tissue contains a strong band at 1006 cm^{-1}, usually attributed to phenylalanine, while normal tissue does not. The ratios of the 1303 to 1265 cm^{-1} bands, indicative of the ratio of lipid to protein content, are very different between the tissue types, and the amide I band centered around 1656 cm^{-1} is much wider in the tumor as compared to normal—again indicative of increased relative protein contributions in the cancerous tissues. Also, the 1445 cm^{-1} CH stretch band is relatively more intense in normal tissue, and the normal tissue contains a carbonyl stretch peak around 1750 cm^{-1}, typically due to fat content, while the tumor tissue does not.

Figures 6 and 7 show H&E stained tissue sections and the SORS spectra from those sections from three different [Figs. 6, 7(a) and 7(b), and 7(c) and 7(d)] *in vitro* tumor samples. In all histological images, the "*S*" arrow indicates the placement of the source fiber, while the "R1," "R2," etc., labels denote the location of the individual collection fiber rings. In the tissue sample from Fig. 6(a), the probe was delivering light to a large fatty area, as seen by the whitish (formerly lipid-filled) vacuoles, while only the outermost collection fibers were placed over a portion of the tumor, which comprises the remainder of the darkly stained section. Since spectral differences among detector rings in Fig. 6(b) are visually subtle, except for differences around 1445 cm^{-1}, close-ups of the three spectral regions are shown in Figs. 6(c)–6(e). These plots show definite trends indicating that the closer rings are sampling normal tissue, while the outer rings are picking up slight spectral contributions from the tumor as well. By qualitatively comparing the spectra in Fig. 6(b) with the typical pure normal and tumor spectra from Fig. 5, these trends include the increasing presence of the 1006 cm^{-1} peak, the lesser relative contributions from the 1303 and 1445 cm^{-1} peaks, and the increasing width of the 1656 cm^{-1} peak as the source-detector offset increases. These trends are

Fig. 4 Mean (n = 12) SNR, plus or minus one standard deviation, for spectra of chicken muscle binned within each detector ring and normalized to SNR of first ring.
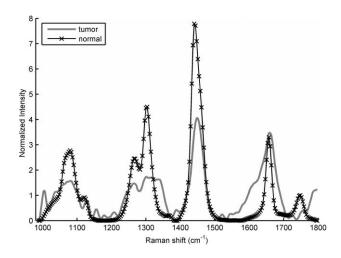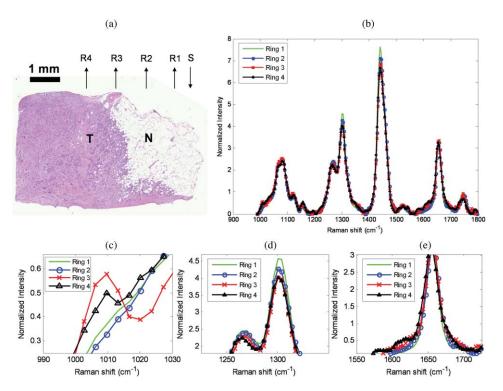
Fig. 6 (a) H&E stained tissue section of IDC sample with a large area of normal fat (white colored area with "N") on the right, and solid IDC tumor (darkly stained area with "T") on the left. Arrows indicate the placement of the source fiber (S) and each of the detector rings. (b) Binned SORS spectra for each detector ring from tissue in *A*. To aid in the visualization of relevant, but subtle spectral changes, zoomed-in versions are shown for (c) the 1006 cm^{-1} phenylalanine peak, (d) the 1265 cm^{-1} amide III and 1303 cm^{-1} lipid peak, and (e) the shoulders of the 1656 cm^{-1} amide I peak.

similar to those seen in the earlier report of SORS on layered breast tissues,[16] but in this case, the tissue boundary was vertical rather than horizontal.

The example in Figs. 7(a) and 7(b) provides an illustration of what happens with smaller layers of normal tissue over a tumor. Figure 7(a) shows a sample with a large tumor region, but with pockets of normal adipose cells near the surface, including directly under the location where the excitation fiber from the probe was placed. From Fig. 7(b), in comparison to Fig. 5, the spectrum from the smallest S-D offset mostly contains features indicative of normal fatty breast tissue, while spectra from the larger S-D offsets contain features indicative of tumor spectral signatures, as noted above. The sample from Figs. 7(c) and 7(d) is included to confirm that if the excitation side of the probe is placed on the tumor tissue overlying normal tissue (i.e., the opposite of margin analysis), then the inner detector rings picked up tumor signatures, while the outer rings picked up the appropriate degree of normal spectral signatures. Thus, it is clear that the different detector rings are sampling different volumes, as desired.

To simplify the "margin analysis" procedure, the spectra from each detector ring were averaged to create one composite spectrum per *in vitro* sample. Thus, a single histological classification could be correlated to a single spectral classification. Table 1 shows the confusion matrix for the classification of these composite spectra with SMLR. This analysis showed an excellent ability for SORS to evaluate margin status in breast specimens, with 95% sensitivity and 100% specificity, and an area under the receiver operating characteristic curve of 0.993. Alternatively, the discrimination was performed with

94% negative predictive value (NPV) and 100% positive predictive value (PPV). The one false negative came from a tumor sample which, after formalin fixation and sectioning, was found to have a ~1.5 mm layer of normal tissue between the measurement site and the tumor. Since it has been shown that normal tissue margins tend to shrink by an average of 33% during formalin fixation,[36] it is possible that this normal layer was at least 2-mm thick when the spectra were obtained.

4 Discussion

This manuscript presents the design, testing, and implementation of a multiseparation SORS probe for use in evaluating surgical margin status following partial mastectomies. The design shown in Fig. 3 was based on the results from our earlier experimental and simulation-based studies,[16,25] and from the SNR simulation results from Fig. 2. To ensure that the SNRs were comparable across the different detector rings, a series of measurements was performed using the common soft tissue optical phantom of a chicken breast. As seen in Fig. 4, the design of adding an additional collection fiber for each further-offset ring worked well to keep the SNR of each ring no more than ~30% different from the others. Given the exponential shape of Fig. 2, it would be very difficult to design a probe to both sample the desired depths in tissue and achieve even better equilibration of SNR among the various detector rings. Besides the SNR balancing, the probe design from Fig. 3 also appeared to sample tissue to the expected depths based on earlier experimental[16] and simulation[25] results. This conclusion is supported by the
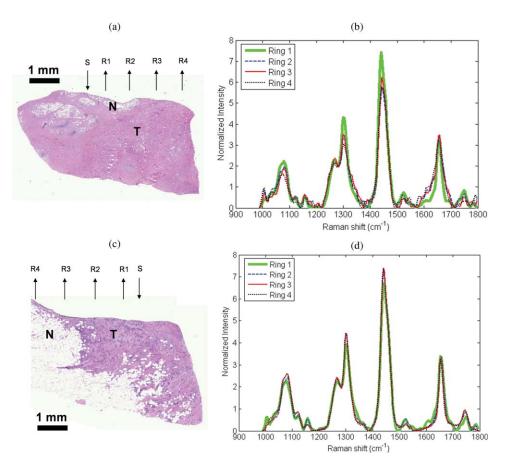
Fig. 7 (a) H&E stained tissue section of ILC sample with pockets of normal fat ("N") near surface of otherwise darkly stained tumor tissue ("T"). Arrows again represent placements of fibers. (b) SORS spectra for each detector ring from tissue in *A*. (c) H&E stained tissue section of IDC sample with underlying fat and (d) corresponding SORS spectra.

success shown in Table 1 for classifying spectra according to the margin status using 2 mm as the cutoff value for negative versus positive classification.

The ability of the detector rings to sample different volumes is demonstrated in Figs. 6 and 7. From Fig. 6(a), it is clear that the SORS probe was placed over two very different regions of tissue for that specimen. A large area of normal fatty tissue was found directly under the excitation fiber and the first 2 to 3 detector rings, while the outermost 1 to 2 detector rings were placed against the tumor. Comparing Figs. 6(b)–6(e) to the pure normal and tumor spectra from Fig. 5, rings 1 and 2 show essentially no tumor spectral signatures. Given this, a standard Raman probe placed in the same spot would not detect any positive margin

Table 1 Confusion matrix for "margin analysis" on *in vitro* specimens.

		Spectral margin status		
		Negative	Positive	
Histopathology	Negative	15	0	Specificity: 100%
Margin Status	Positive	1	19	Sensitivity: 95%
		NPV: 94%	PPV: 100%	

findings at this point. The 3rd and 4th rings of the probe were able to detect slight tumor contributions though, indicating that they successfully sampled a different volume of tissue than the inner rings. While most spectral regions showed an increasing tumor contribution from ring 1 to ring 4, ring 3 had a stronger relative 1006 cm^{-1} peak than ring 4. Possible causes for this include a slight misalignment between the probe and histology, especially considering the rotation of the probe, or inconsistent biochemical composition of the tumor tissue sampled by rings 3 and 4. A similar situation to Fig. 6 is seen in Figs. 7(a) and 7(b), although only the first detector ring was sensitive to a small (< 1 mm thick) fat layer on the surface, while the outer rings sampled deeper and more radially distal tissue volumes. It should be noted that in the fixation of samples, the fat regions tend to shrink,[36] so the measurement surface of these specimens were likely flatter during signal acquisition. Also, all specimens were cut after fixation and before sectioning to make a given section contain only the interrogated tissue region, so the fibers were never placed over the edge of any sample.

The opposite situation of the above two samples is shown in Figs. 7(c) and 7(d), where the source fiber was placed over a tumor region ∼1 to 1.5 mm thick, with normal tissue underneath; outer detector rings were placed over a considerably thinner tumor layer with more underlying normal tissue. Taken with the above results, these panels demonstrate that the spectral signatures collected with the SORS probe vary appropriately as

a function of S-D offset according to tissue type and location, not via any systematic response. Since some normal Raman signatures are present even in ring 1 spectra, Figs. 7(c) and 7(d) also shows that the presence of normal tissue under tumor tissue would be easier to detect than the current problem, since fat is a stronger Raman scatterer than tumor tissue.[25]

Given these findings regarding sampling depths and volumes, the composite spectra were used for margin analysis on intact breast specimens in the laboratory. Since the SNR is approximately equal in all 4 rings (see Fig. 4), averaging them provides information about the entire sampling volume in a single spectrum. This method also simplifies the analysis procedure; if spectra from individual rings were used, it would be difficult to determine how to correlate certain ones with pathology findings. For example, although all spectra in Figs. 6(b) and 7(b) were from tissue sites that would be deemed positive margins within the spatial extent of the probe, it is unlikely that the innermost rings were actually detecting any signal from tumor tissues. A possible approach for using the individual spectra would be to label a measurement site positive if any spectrum from the 4 rings is predicted to be from a positive margin, but the aforementioned correlation issue arises in the training of such an algorithm for a retrospective analysis. Many normal looking spectra, like ring 1 from Fig. 6(b), would be labeled as a tumor and would likely cause difficulties for discrimination algorithms trying to create decision boundaries between negative and positive margins.

A binary diagnostic algorithm simpler than SMLR may seem like a more appropriate approach in this analysis, but the SMLR algorithm was able to significantly reduce the dimensionality of the data from the initial size of 232 variables (one per 3.5 cm^{-1} bin) to perform its classification. In addition, SMLR provides a probability of class membership that would be very useful in a clinical application. A surgeon could act differently if the probability of a margin being negative is 99% versus 51%, although in either case, the diagnosis would be negative.

The results from using SMLR to classify the composite SORS spectra according to margin status are shown in Table 1. With only one false negative, the sensitivity, specificity, NPV, and PPV were all at least 94%. For this clinical application, perhaps the most important variable for long term studies is NPV, since a surgeon needs to be confident in any diagnosis of negative margin status to prevent recurrence of the disease or unnecessary second operations. For the single false negative result in this study, the normal layer overlying the tumor was found to be ~1.5 mm thick upon histological examination, but prior to formalin fixation, this layer was likely around or slightly greater than 2-mm thick,[36] which would surpass the sampling capabilities of the SORS probe. It may also be possible that a slight misalignment between the probe and the point of the histological section led to an error in the margin size determination. In addition, there is not a universal standard among hospitals of a minimum margin size required during breast conserving surgery; rather, some locations use 2 mm, some use 1 mm, and others simply require that no cancer cells be found on the surface of the specimen.[3] We used 2 mm as the cutoff in this study because that value provides the best prognosis for patients[3] and is the most stringent standard for proving the value of SORS.

The classification results above compare extremely favorably with current intraoperative margin evaluation techniques.[5–8] For example, the reported sensitivity of "touch prep" is as low as 8% (Ref. 5); a simple visual examination has sensitivity and specificity of approximately 50% and 72%, respectively;[7] and frozen section pathology, though its sensitivity and specificity per slide are generally > 90%, suffers from a sampling error that brings per-specimen classification accuracy (i.e., overall designation of whether a second operation is required) below 85% (Ref. 8). Another optical approach for intraoperative margin evaluation is to image an entire margin (i.e., one of the six facets of the "cuboidal" excised specimen) at once with autofluorescence and/or diffuse reflectance modalities. While our group has published a small study on the topic,[13] it has been extensively researched in recent years by the Ramanujam group.[9–12] Using extracted optical properties of the tissue from visible diffuse reflectance, that group has achieved 79% sensitivity and 67% specificity for discriminating normal from positive or close (< 2 mm) margins for a set of 48 patients.[11] Better classification—100% sensitivity and 82% specificity—was achieved by Nguyen et al. using OCT,[14] though the sample size of 20 patients was much smaller, and the technology in its current state still relies on a subjective analysis of the images. The biggest challenge for the SORS approach presented here, especially compared with the above two techniques, is adapting the probe and other system components to interrogate larger areas of tissue in a shorter time.

The various optical approaches to an intraoperative margin evaluation all hold significant potential for improving the standard of care, though each currently has its own strengths and weaknesses. To date, no method has demonstrated the combination of sampling speed, volume, and diagnostic accuracy needed for widespread clinical implementation. The initial work presented here has demonstrated the feasibility and promise of using SORS to evaluate the margin status on intact breast specimens in a laboratory setting. Studies are currently underway on using the same approach in a clinical setting, and initial results are equally as promising as the laboratory measurements. These clinical SORS measurements for breast tumor surgical margin evaluation will be the subject of future manuscripts.

References

1. American Cancer Society, "Cancer Facts and Figures 2010," American Cancer Society, Atlanta (2010).
2. K. I. Bland and E. M. Copeland III, Eds., *The Breast: Comprehensive Management of Benign and Malignant Disorders*, Saunders, St. Louis (2004).
3. A. Taghian, M. Mohiuddin, R. Jagsi, S. Goldberg, E. Ceilley, and S. Powell, "Current perceptions regarding surgical margin status after breast-conserving therapy: results of a survey," *Ann. Surg.* **241**(4), 629–639 (2005).
4. K. C. Horst, M. C. Smitt, D. R. Goffinet, and R. W. Carlson, "Predictors of local recurrence after breast-conservation therapy," *Clin. Breast Cancer* **5**(6), 425–438 (2005).

5. E. K. Valdes, S. K. Boolbol, I. Ali, S. M. Feldman, and J. M. Cohen, "Intraoperative touch preparation cytology for margin assessment in breast-conservation surgery: does it work for lobular carcinoma?," *Ann. Surg. Oncol.* **14**(10), 2940–2945 (2007).

6. T. P. Olson, J. Harter, A. Munoz, D. M. Mahvi, and T. Breslin, "Frozen section analysis for intraoperative margin assessment during breast-conserving surgery results in low rates of re-excision and local recurrence," *Ann. Surg. Oncol.* **14**(10), 2953–2960 (2007).

7. G. C. Balch, S. K. Mithani, J. F. Simpson, and M. C. Kelley, "Accuracy of intraoperative gross examination of surgical margin status in women undergoing partial mastectomy for breast malignancy," *Am. Surg.* **71**(1), 22–27 (2005).

8. J. C. Cendan, D. Coco, and E. M. Copeland, III, "Accuracy of intraoperative frozen-section analysis of breast cancer lumpectomy-bed margins," *J. Am. Coll. Surg.* **201**(2), 194–198 (2005).

9. T. M. Bydlon, S. A. Kennedy, L. M. Richards, J. Q. Brown, B. Yu, M. K. Junker, J. Gallagher, J. Geradts, L. G. Wilke, and N. Ramanujam, "Performance metrics of an optical spectral imaging system for intraoperative assessment of breast tumor margins," *Opt. Express* **18**(8), 8058–8076 (2010).

10. J. Q. Brown, T. M. Bydlon, L. M. Richards, Y. Bing, S. A. Kennedy, J. Geradts, L. G. Wilke, M. K. Junker, J. Gallagher, W. T. Barry, and N. Ramanujam, "Optical assesssment of tumor resection margins in the breast," *IEEE J. Sel. Top. Quantum Electron.* **16**(3), 530–544 (2010).

11. L. G. Wilke, J. Q. Brown, T. M. Bydlon, S. A. Kennedy, L. M. Richards, M. K. Junker, J. Gallagher, W. T. Barry, J. Geradts, and N. Ramanujam, "Rapid noninvasive optical imaging of tissue composition in breast tumor margins," *Am. J. Surg.* **198**(4), 566–574 (2009).

12. S. Kennedy, J. Geradts, T. Bydlon, J. Q. Brown, J. Gallagher, M. Junker, W. Barry, N. Ramanujam, and L. Wilke, "Optical breast cancer margin assessment: an observational study of the effects of tissue heterogeneity on optical contrast," *Breast Cancer Res.* **12**(6), R91 (2010).

13. M. D. Keller, S. K. Majumder, M. C. Kelley, I. M. Meszoely, F. I. Boulos, G. M. Olivares, and A. Mahadevan-Jansen, "Autofluorescence and diffuse reflectance spectroscopy and spectral imaging for breast surgical margin analysis," *Lasers Surg. Med.* **42**(1), 15–23 (2010).

14. F. Nguyen, A. Zysk, E. Chaney, J. Kotynek, U. Oliphant, F. Bellafiore, K. Rowland, P. Johnson, and S. Boppart, "Intraoperative evaluation of breast tumor margins with optical coherence tomography," *Cancer Res.* **69**, 8790–8796 (2009).

15. A. S. Haka, Z. Volynskaya, J. A. Gardecki, J. Nazemi, J. Lyons, D. Hicks, M. Fitzmaurice, R. R. Dasari, J. P. Crowe, and M. S. Feld, "*In vivo* margin assessment during partial mastectomy breast surgery using raman spectroscopy," *Cancer Res.* **66**(6), 3317–3322 (2006).

16. M. D. Keller, S. K. Majumder, and A. Mahadevan-Jansen, "Spatially offset Raman spectroscopy of layered soft tissues," *Opt. Lett.* **34**(7), 926–928 (2009).

17. M. V. Schulmerich, J. H. Cole, J. M. Kreider, F. Esmonde-White, K. A. Dooley, S. A. Goldstein, and M. D. Morris, "Transcutaneous Raman spectroscopy of murine bone *in vivo*," *Appl. Spectrosc.* **63**(3), 286–295 (2009).

18. N. A. Macleod and P. Matousek, "Deep noninvasive Raman spectroscopy of turbid media," *Appl. Spectrosc.* **62**(11), 291A–304A (2008).

19. N. A. Macleod, A. Goodship, A. W. Parker, and P. Matousek, "Prediction of sublayer depth in turbid media using spatially offset Raman spectroscopy," *Anal. Chem.* **80**(21), 8146–8152 (2008).

20. N. Stone, R. Baker, K. Rogers, A. W. Parker, and P. Matousek, "Subsurface probing of calcifications with spatially offset Raman spectroscopy (SORS): future possibilities for the diagnosis of breast cancer," *Analyst.* **132**(9), 899–905 (2007).

21. M. V. Schulmerich, K. A. Dooley, T. M. Vanasse, S. A. Goldstein, and M. D. Morris, "Subsurface and transcutaneous Raman spectroscopy and mapping using concentric illumination rings and collection with a circular fiber-optic array," *Appl. Spectrosc.* **61**(7), 671–678 (2007).

22. P. Matousek, "Deep noninvasive Raman spectroscopy of living tissue and powders," *Chem. Soc. Rev.* **36**(8), 1292–1304 (2007).

23. M. V. Schulmerich, K. A. Dooley, M. D. Morris, T. M. Vanasse, and S. A. Goldstein, "Transcutaneous fiber optic Raman spectroscopy of bone using annular illumination and a circular array of collection fibers," *J. Biomed. Opt.* **11**(6), 060502 (2006).

24. P. Matousek, E. R. Draper, A. E. Goodship, I. P. Clark, K. L. Ronayne, and A. W. Parker, "Noninvasive Raman spectroscopy of human tissue *in vivo*," *Appl. Spectrosc.* **60**(7), 758–763 (2006).

25. M. D. Keller, R. H. Wilson, M. A. Mycek, and A. Mahadevan-Jansen, "Monte Carlo model of spatially offset Raman spectroscopy for breast tumor margin analysis," *Appl. Spectrosc.* **64**(6), 607–614 (2010).

26. K. Vishwanath and M. A. Mycek, "Time-resolved photon migration in bi-layered tissue models," *Opt. Express* **13**(19), 7466–7482 (2005).

27. K. Vishwanath and M. A. Mycek, "Do fluorescence decays remitted from tissues accurately reflect intrinsic fluorophore lifetimes?," *Opt. Lett.* **29**(13), 1512–1514 (2004).

28. K. Vishwanath, B. Pogue, and M. A. Mycek, "Quantitative fluorescence lifetime spectroscopy in turbid media: comparison of theoretical, experimental, and computational methods," *Phys. Med. Biol.* **47**(18), 3387–3405 (2002).

29. C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological Raman spectra," *Appl. Spectrosc.* **57**(11), 1363–1367 (2003).

30. A. Mahadevan-Jansen and R. Richards-Kortum, "Raman spectroscopy for the detection of cancers and precancers," *J. Biomed. Opt.* **1**, 31–70 (1996).

31. C. Reble, I. Gersonde, C. A. Lieber, and J. Helfmann, "Influence of tissue absorption and scattering on the depth dependent sensitivity of Raman fiber probes investigated by Monte Carlo simulations," *Biomed. Opt. Express* **2**(3), 520–533 (2011).

32. C. M. Krishna, J. Kurien, S. Mathew, L. Rao, K. Maheedhar, K. K. Kumar, and M. Chowdary, "Raman spectroscopy of breast tissues," *Expert Rev. Mol. Diagn.* **8**(2), 149–166 (2008).

33. A. S. Haka, Z. Volynskaya, J. A. Gardecki, J. Nazemi, R. Shenk, N. Wang, R. R. Dasari, M. Fitzmaurice, and M. S. Feld, "Diagnosing breast cancer using Raman spectroscopy: prospective analysis," *J. Biomed. Opt.* **14**(5), 054023 (2009).

34. B. Krishnapuram, L. Carin, M. A. T. Figueiredo, and A. J. Hartemink, "Sparse multinomial logistic regression: Fast algorithms and generalization bounds," *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(6), 957–968 (2005).

35. S. K. Majumder, M. D. Keller, M. C. Kelley, F. I. Boulos, and A. Mahadevan-Jansen, "Comparison of autofluorescence, diffuse reflectance, and Raman spectroscopy for breast tissue discrimination.," *J. Biomed. Opt.* **13**(5), 054009 (2008).

36. B. H. Yeap, S. Muniandy, S. K. Lee, S. Sabaratnam, and M. Singh, "Specimen shrinkage and its influence on margin assessment in breast cancer," *Asian J. Surgery* **30**(3), 183–187 (2007).