Intrinsic near-infrared spectroscopic markers of breast tumors

Shwayta Kukreti¹,²,⁴, Albert Cerussi¹, Bruce Tromberg¹,³, and Enrico Gratton¹,²,³,⁴,*
¹ Beckman Laser Institute, University of California, Irvine CA 92612
² Laboratory for Fluorescence Dynamics, University of California, Irvine CA 92612
³ Department of Biomedical Engineering, University of California, Irvine CA 92612
⁴ Biophysics and Computational Biology Program, University of Illinois, Urbana-Champaign, IL 61801

Abstract

We have discovered quantitative optical biomarkers unique to cancer by developing a double-differential spectroscopic analysis method for near-infrared (NIR, 650-1000nm) spectra acquired non-invasively from breast tumors. These biomarkers are characterized by specific NIR absorption bands. The double-differential method removes patient specific variations in molecular composition which are not related to cancer, and reveals these specific cancer biomarkers. Based on the spectral regions of absorption, we identify these biomarkers with lipids that are present in tumors either in different abundance than in the normal breast or new lipid components that are generated by tumor metabolism. Furthermore, the O-H overtone regions (980-1000 nm) show distinct variations in the tumor as compared to the normal breast. To quantify spectral variation in the absorption bands, we constructed the Specific Tumor Component (STC) index. In a pilot study of 12 cancer patients we found 100% sensitivity and 100% specificity for lesion identification. The STC index, combined with other previously described tissue optical indices, further improves the diagnostic power of NIR for breast cancer detection.

Keywords

Breast Cancer; Optical Spectra; Near Infrared Spectroscopy (NIRS) Lipids; Lipid Spectroscopy

1. Introduction

Do intrinsic quantitative optical biomarkers exist for malignant tumors that can be measured non-invasively? If this will be found, then optical biopsy using non-invasive methods could be performed with great advantages in terms of cost, number of procedures, speed and availability upon existing methods. Optical spectroscopy is not invasive and offers a unique view of tissue because it is sensitive to compositional and functional characteristics. Spectroscopic methods have employed absorption, scattering, and fluorescent contrast (endogenous and exogenous) to detect tumors in breast, cervix, skin, and esophagus over wavelengths ranging from the ultraviolet through the infrared. Depending upon the optical wavelengths, light can interrogate superficial to cm-thick tissues.[1] Although the penetration of light in tissues is strongly wavelength dependent, light may be delivered to most tissues either non-invasively (topical application) or by the use of intra-operative or minimally invasive probes (i.e., endoscopy). Our contribution will deal exclusively with breast tissues

*Corresponding author, Enrico Gratton, University of California-Irvine, 3210 Natural Sciences II Building, Irvine, CA 92697-2715, USA, Phone: 949-824-2674, Fax: 949-824-1727, Email: egratton22@yahoo.com.
exposed to NIR light, but the general principles of our method could be applied to any cancer type.

Within the NIR spectral region (650-1000 nm), light penetrates deeply enough to transilluminate cm-thick tissues such as the breast. NIR is sensitive to the four major absorbing components in breast tissues: oxyhemoglobin, deoxyhemoglobin, water, and bulk lipids.[2,3]

The main hypothesis of NIR cancer detection has been that malignant transformation changes the balance of hemoglobin (oxygenation state), water, and bulk lipids in tissue. Differences in optical signatures between tissues are manifestations of multiple physiological changes associated with factors such as vascularization, cellularity, oxygen consumption, edema, fibrosis, and remodeling. [8] Several groups have employed NIR optical methods to measure subtle physiological alterations in healthy breast tissue,[7,9-12] cancerous breast tissues,[8, 13-20] and malignant tumors treated with neoadjuvant chemotherapy.[21-23]

The transport of NIR light in tissues is dominated by multiple-scattering, which complicates the quantitative recovery of the amounts of these four components from the absorption spectra from tissues. Light transport can be modeled as a diffusive process where photons behave as stochastic particles much like the bulk movement of molecules or heat. The most accurate tissue optical absorption spectra can only be obtained by separating light absorption from scattering, which can be performed using a model-based approach alongside time- or frequency-domain measurements.[4-6] These absorption spectra are then translated into hemoglobin, water, and bulk lipids, using the Beer-Lambert law along with known component extinction coefficients (Figure 1) (for a detailed description see [7]).

After many years of research it is realized that although changes in NIR measured components are altered significantly by tumors, they are not specific for cancer; the same components are found in both normal and malignant tissues. Thus the detection of cancer is based upon crossing a threshold in hemoglobin, and sometimes water, and bulk lipids. Complications can arise in separating malignant from benign lesions, since benign lesions are also known to change water, hemoglobin, and bulk lipids in a similar fashion to malignant lesions.[7,19] Exogenous contrast agents have been developed for both absorption and fluorescence contrast, in order to increase sensitivity and specificity.[24]

However, it is known that tumors alter the composition of tissues in other ways that are not accounted for in traditional NIR spectroscopic models. Alterations in protein expression, lipid and water states, and the presence of hemoglobin breakdown products can occur in malignant transformation.[25-28] These biochemical state alterations, or additional biochemical components, are likely to be small contributions to the overall tumor absorption spectrum. Even if smaller, these contributions should be unique to tumor tissues, and provide an endogenous spectroscopic signature that is specific of the lesion. To summarize: our hypothesis is that the optical absorption spectra of malignant tissues differ from normal spectra not only due to bulk absorber concentration differences, but also due to small changes of as yet unknown components.

In order to discern these subtle alterations in tumor absorption spectra that are not accounted for in traditional NIR component models of breast tissues, we have developed a double-differential method. The method nulls the effect of the four basis components to reveal the small spectral changes unique to cancer. Note that this method only works because we have two independent controls as will be discussed later. The results of our analysis will show only intrinsic, unique quantitative biomarkers of cancers which are not included in the basis spectra of the four components (hemoglobin, water and bulk lipids). We hypothesize the biochemical/physical origin of these spectral bands to reside in tumor specific changes in types and abundance of lipid components and changes in the O-H overtone region. Future work will
determine the chemical origin of these spectral changes and improve breast cancer diagnostics by using these specific endogenous molecular markers to improve understanding of molecular changes in cancerous tissue.

2. Materials and Methods

A. Instrumentation

The detection and quantitation of subtle spectral shifts to fingerprint disease requires accurate reconstruction of tissue absorption spectra uncontaminated by the effects of scattering. For this purpose we have employed a technique known as Diffuse Optical Spectroscopy (DOS), more specifically we use the well established method of Steady State Frequency Domain Photon Migration (SSFDPM) which can quantitatively measure scattering and absorption independently from each other. [29-31] The particular instrument used was the Laser Breast Scanner (LBS), which is a bedside-capable instrument that has been used in several clinical trials for breast cancer detection [32,33] and monitoring response to neoadjuvant chemotherapy.[21,22] The LBS employs a handpiece similar to an ultrasound probe which was placed on the surface of the breast to recover tissue optical spectra from 650 to 1000 nm. Typically DOS samples a low number of spatial locations with a large spectral bandwidth.

B. Data Processing

Data were analyzed in three stages: (1) calculation of absorption and scattering spectra, (2) calculation of absorber concentrations, and (3) calculation of unique tumor spectra by the double differential method. Stages 1 & 2 have been detailed in the literature and are only summarized here.[32] A diffusive model constructed within the semi-infinite geometry is used to translate measured optical signals into absorption and scattering coefficients for each measured wavelength. Using independently measured molar extinction coefficients, the concentrations of the dominant absorbers found in breast tissue can be obtained: oxygenated [O$_2$Hb] and deoxygenated hemoglobin [HHb], water [H$_2$O], and bulk lipids. (Figure 1). The concentrations of these chromophores are then used to diagnose the presence of disease; this approach was outlined in a previous review article in this journal.[34] All data processing was performed using custom software designed in the MATLAB platform.

The unique spectral components were calculated using a double-differential technique applied to the measured absorption spectra, which is described in the Results and Discussion section. The absorbance data were analyzed by the Elantest program (available at ftp://ftp.lfd.uci.edu/lfd/egratton/elantest/elantest.exe

C. Measurement Procedure

We employed a measurement procedure which has been detailed in the literature.[32] Briefly, optical linescans were generated by moving the handheld probe to a set of discrete positions on the breast surface in 10 mm steps (Figure 2a). Tumor locations were known a priori from mammography, ultrasound, and palpation. Absorption and reduced scattering spectra were measured at each grid location (Figure 2b). A complete measurement of tissue absorption and reduced scattering spectra required approximately 25 seconds at each linescan location. Linescans were repeated twice at each grid location to evaluate placement variations. The fraction of tumor to normal tissue sampled by the light depended upon the tissue optical properties and the lesion depth.

D. Patient Characteristics

All subjects were informed and provided written informed consent to participate in the studies under a protocol approved by the Internal Review Board of the University of California at
Ultrasound and surgical pathology reports were utilized to determine type, localization, and tumor lesion extent. In this pilot study patients ranged in age from 32-57 years, 6 pre- and 6 post-menopausal, with pathologically confirmed diagnosis of invasive ductal carcinoma with one patient diagnosed with adenocarcinoma with lobular features.

3. Results and Discussion

A. Breast tissue spectra

Typical NIR diffuse optical spectra of normal and tumor-containing tissues are provided for a Patient 30, a 45-year old pre-menopausal patient with a 29 mm invasive ductal carcinoma in the right breast.

In Figure 3 we show representative spectra for the measured wavelength dependence of scattering and absorption from 650 to 990 nm. These spectra correspond to a single spatial location over the tumor, and the equivalent tissue region on the contra lateral breast which is considered normal, having no known lesion. Repeat measurements over the years of data acquisition have shown that the spectral shape of tumor and normal tissue is well conserved. Variation in the magnitude of the absorption values due to probe handling is on the order of 3-5%. [32] The wavelength dependence of the tissue scattering (i.e., the scatter power, SP) differs substantially between tumor and normal tissue suggesting differences in the density and size of scattering centers between tumor and normal tissue regions (Figure 2a).[32,35,36]

The measured NIR spectra demonstrate distinctive differences between tumor and normal tissues. Briefly, the spectra over tumor-containing tissue exhibits higher overall absorption along with notable spectral differences in the region above 900 nm. The distinctive peak in the normal absorption spectrum at 930 nm is representative of the vibrational overtones of lipid C-H bonds. Higher lipid absorption is typically seen in normal and postmenopausal breast tissue as compared to tumor-containing breast tissue and pre-menopausal breast tissue, respectively. This prominent lipid peak results from the higher concentration of bulk lipid in normal adipose tissues relative to tumor tissues.[32]

By applying the results of the four component basis spectra fit we determined concentrations for these components, and we calculated the Tissue Optical Index (TOI), which has been discussed in a previous manuscript in this journal[34]. Table 1.

\[
\text{TOI} = \frac{\text{Deoxyhemoglobin}}{\text{Water}} \times \frac{\text{Water}}{\text{BlukLipid}} 
\]

Briefly the TOI is a contrast function empirically designed using the traditional components (recovered concentrations of deoxyhemoglobin, water, and bulk lipids) to accentuate the differences between tumor and normal tissue. While the TOI has been shown to identify some types of tumors, the TOI index is not highly specific. In fact for our set of 12 patients the TOI Index showed a sensitivity of 75% and specificity of 75% based on a threshold of 3.5 that gives three false positives and three false negatives. This makes differential diagnosis solely based on the four traditional components problematic. Thus, using this approach, we had no unique “marker” or “signature” of cancer.

B. Double differential method

In the double-differential method we want to determine if there are other spectral differences that cannot be accounted for by the different amounts of the four basis components. We first
calculate the average spectrum of the normal breast. Then we calculate the differences between this average spectrum and the spectrum at each location (including the normal breast). If the only components present are the ones included in the four basis spectra, then the difference spectrum (at each location) could be completely fitted by the using these four components. The coefficient of the fit will provide us with the different amounts of the four components at each location (the same as for the TOI described in the previous paragraph). However, if the fit is not perfect, the residual of the fit (if highly correlated) will provide the additional spectra which are not included in the four basis components. This residual is the Specific Tumor Component (STC).

The double-differential method has two major advantages over the conventional spectroscopic modeling approach. Firstly, the method uses each patient as their own control, so that variations in the “known” extinction spectra are effectively cancelled out. Thus the double-differential method is insensitive to intra-subject variations that arise due to the four component spectroscopic model. Secondly, the method accounts for concentration changes within subjects between normal and tumor tissue if the only differences are in the abundance of the four components. The residual spectrum for different locations of the unaffected breast should essentially be a flat line with values close to zero. If the tumor differs from normal tissues only in concentration changes of the four components, again the residual spectrum should be essentially zero. However, if spectral components are not accounted for in the four component fit, they should appear only in the tumor regions.

C. Double differential method applied to breast spectra

Here we outline the double-differential method applied to Patient 30, the same patient presented in Figure 3. In Figure 4a we show residual spectra for all of the positions on the normal breast. In Figure 4b we show residual spectra over a tumor containing position and positions over normal tissue surrounding the lesion in the tumor-containing breast.

As stated earlier, we have two internal controls, contra lateral normal breast tissue and the normal breast tissue surrounding the lesion on the tumor containing breast. Note that all of the residual spectra over normal breast tissue are essentially featureless and provide random values around zero. This control is important in that it proves that variations in normal breast tissue are due only to the natural compositional differences from oxyhemoglobin, deoxyhemoglobin, water and bulk lipids.

D. STC spectra in 12 cancer patients

In Figure 5, we show the STC spectra for 12 cancer patients of this pilot study at the location displaying the largest variation. These residual spectra have small amplitudes (about 1% of original spectra). Clearly, the STC spectra are highly reproducible across all patients in this pilot study. There are patient dependent variations in several wavelength regions (Figure 5), however all tumor STC spectra display a similar spectral shape. Figure 6 displays a comparison of the average of the 12 tumor spectra presented in Figure 5, along with average of normal spectra obtained from the equivalent position on the unaffected breast for each patient. From Figure 6 it is clear that these residual spectra are not random, unlike the residuals for the normal breast tissue which are featureless.

From inspection of the average STC spectra, we find there are roughly 5 spectral regions where systematic differences are observed. We have labeled these to be Region 1: 650-665nm; Region 2: 730-800nm; Region 3: 875-930nm; Region 4: 930-960nm; and Region 5: 980-990nm. (Figure 6). To quantify the amount of STC spectra, which should indicate the amount of biomarker, we calculate the local residual variance for each spectral region, defined by:
The local variance, $L_k$, is a function of position on the breast given by $x$ and $y$ coordinates. The index $k$ indicates a given spectral region and $N_k$ indicates the total number of wavelengths in a particular region. $STC_i(\lambda_i, x, y)$ is the value of the STC spectra at a given wavelength. The STC index is the sum of all local variance $L_k$. Across the 12 patients in this pilot study, the STC index displays the maximum value over the tumor regions. This STC index value can be significantly greater than the value over the surrounding normal tissue as well as the normal contra lateral breast tissue (Table 1).

For this pilot study of 12 patients we were able to obtain 100% sensitivity and 100% specificity using the STC index. We set the threshold above 51.5 which separates the two groups. We emphasize that these sensitivity and specificity values are only provided for general reference. Data on a larger subject pool will be reported in another manuscript, thereby providing more significant values. We note that the double-differential method assumes that normal and diseased tissues coexist in the same patient. Furthermore, this method is not required that a priori knowledge of the lesion location be available.

E. Origins of STC

We are investigating the biochemical/physical origin of the STC spectral signature, which will be the subject of a future manuscript. We believe Region 1 (650-665nm) is mainly affected by the changes in hemoglobin absorption. Perhaps there may be other types of hemoglobins or breakdown products in tumors. With regards to Region 2 (730-800nm), we see a distinct negative peak in the STC spectra. According to the definition of STC, a negative peak indicates absence of a component or a broadening of the NIR band. Changes in this region may be indicative of lipid metabolism changes, as lipids absorb in this region (Figure 1). In Region 3 (875-930nm) and Region 4 (930-960nm), we observe a negative peak immediately followed by a positive peak. This may be indicative of a spectral shift toward longer wavelengths. This is also the spectral region in which lipids have characteristic absorption spectra (Figure 1). Thus if there are differences in lipid composition in the tumor with respect to the normal tissue, this spectral region is likely to be affected. In Region 5 (980-1000nm) we only have a few points available. However we do observe changes which may be due to differences in the O-H overtones. In this region there are two possible candidates: the O-H of water and lipid oxidation products.

4. Conclusion

We have developed a double-differential method to analyze the near infrared absorption spectra of breast tumors. In this method we consider only the spectral differences between normal and diseased tissue by fitting the difference spectra using the basis components spectra and then analyzing the residuals of this fit. This differential approach can be performed by comparison of regions of normal and tumor breast tissue. With this method we show intrinsic spectroscopic markers of breast tumors in the NIR. These are unique signatures related to the biochemical/physical properties of each type of lesion, as the changes in natural tissue components (oxyhemoglobin, deoxyhemoglobin, water and bulk lipid) as well as the individual physiological variation have already been accounted for. Furthermore, these spectral signatures are biomarkers revealing characteristic absorption bands in the lipid region of 760 and 930nm region that were unnoticed before. The O-H overtone band in the 980nm region also shows distinct variations in the tumor region compared to the normal breast. By quantifying this
information, we constructed the STC index which gives 100% sensitivity and 100% specificity for lesion identification for the 12 patients investigated and has the potential to distinguish tumors on the basis of the lipid composition and/or bound water or lipid oxidation products.

In this paper our focus was to describe the discovery of an intrinsic NIR spectroscopic marker of breast cancer. While here we only provide a description of the double-differential method, in a future manuscript we will describe the mathematical operations for obtaining the STC residual spectra beginning with scatter corrected absorption spectra of breast tissue obtained non-invasively using the SSFDPM method. In the future we will also discuss the difference between the residual spectra for benign and malignant lesions, opening up a new approach for differential diagnosis using optical methods. While this discovery was based on a pilot study of 12 cancer patients, we are analyzing STC residual spectra for a larger population, and comparing subjects who either have lesions which are malignant or benign to normal, having no known lesion.

Acknowledgments

This work was made possible by the NIH NCRR biomedical technology center, XPhoton migration grant, the Laser Microbeam and Medical Program (LAMMP, P41RR01192), the NCI network for translational research in optical imaging (U54CA105480) and the California Breast Cancer Research Program. Beckman Laser Institute programmatic support is provided by the AFOSR and the Beckman Foundation. The authors wish to thank Montana Compton, Natasha Shah, Tran Du, Debbi Gordon, and Amanda F. Durkin for their assistance as well as the patients who generously volunteered for these studies.

References


Figure 1.
Basis spectra set used to describe the major tissue components in the breast. Each component is identified in the figure legend.
FIGURE 2a
Figure 2.
(a) Handheld probe from Laser Breast Scanner placed in contact with the breast during data acquisition. (b) Measurement setup. Points across linescan indicate positions measured.
Figure 3.
Representative tumor spectra from Patient 30 at one position on tumor-containing breast; normal spectra from equivalent position on contra lateral breast. (a) Wavelength dependence of scattering for tumor and normal tissues (b) Absorption spectrum for tumor and normal tissues.
Figure 4.
(a) STC spectra of all positions on normal contra lateral breast tissue for Patient 30. (b) STC spectra over tumor region and surrounding normal tissue on both sides along the linescan for Patient 30.
Figure 5.
STC spectra for 12 patients at the position exhibiting the largest variation in comparison to surrounding normal tissue on tumor-containing breast.
Figure 6.
Average STC spectra for 12 patients for tumor at the location exhibiting the largest variation in comparison to surrounding normal tissue on tumor-containing breast, and for normal tissue at the equivalent position on normal contra lateral breast. Standard error of mean bars placed every 20 nm indicate STC spectral variation in the population. Regions where systematic differences have been observed are labeled as Region 1 (650-665nm), Region 2 (730-800nm), Region 3 (875-930nm), Region 4 (930-960nm), and Region 5 (980-1000nm).
<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Menopausal Status</th>
<th>Tumor classification</th>
<th>Size (mm)</th>
<th>STC Tumor</th>
<th>STC Normal</th>
<th>TOI Tumor</th>
<th>TOI Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>PRE</td>
<td>DC</td>
<td>24</td>
<td>120.1</td>
<td>15.0</td>
<td>3.2</td>
<td>1.2</td>
</tr>
<tr>
<td>38</td>
<td>PRE</td>
<td>DC</td>
<td>24</td>
<td>100.2</td>
<td>9.7</td>
<td>4.9</td>
<td>1.4</td>
</tr>
<tr>
<td>50</td>
<td>POST</td>
<td>AC with L.F</td>
<td>54</td>
<td>216.9</td>
<td>4.4</td>
<td>6.2</td>
<td>0.6</td>
</tr>
<tr>
<td>32</td>
<td>PRE</td>
<td>DC</td>
<td>29</td>
<td>422.9</td>
<td>51.5</td>
<td>9.1</td>
<td>3.9</td>
</tr>
<tr>
<td>57</td>
<td>POST</td>
<td>DC w/ L.N.M</td>
<td>32</td>
<td>115.5</td>
<td>20.0</td>
<td>1.2</td>
<td>0.5</td>
</tr>
<tr>
<td>47</td>
<td>POST</td>
<td>DC</td>
<td>17</td>
<td>609.3</td>
<td>8.4</td>
<td>20.0</td>
<td>1.1</td>
</tr>
<tr>
<td>45</td>
<td>PRE</td>
<td>DC</td>
<td>29</td>
<td>162.5</td>
<td>7.8</td>
<td>20.5</td>
<td>1.0</td>
</tr>
<tr>
<td>44</td>
<td>PRE</td>
<td>DC</td>
<td>16</td>
<td>62.7</td>
<td>12.9</td>
<td>12.0</td>
<td>4.0</td>
</tr>
<tr>
<td>49</td>
<td>POST</td>
<td>DC</td>
<td>70</td>
<td>148.3</td>
<td>9.3</td>
<td>3.4</td>
<td>0.8</td>
</tr>
<tr>
<td>41</td>
<td>PRE</td>
<td>DC</td>
<td>40</td>
<td>86.1</td>
<td>4.2</td>
<td>14.4</td>
<td>1.8</td>
</tr>
<tr>
<td>53</td>
<td>POST</td>
<td>DC</td>
<td>24</td>
<td>132.3</td>
<td>6.3</td>
<td>10.2</td>
<td>4.5</td>
</tr>
<tr>
<td>57</td>
<td>POST</td>
<td>DC</td>
<td>31</td>
<td>135.2</td>
<td>4.1</td>
<td>4.5</td>
<td>0.6</td>
</tr>
</tbody>
</table>