Tissue oxygen saturation predicts response to breast cancer neoadjuvant chemotherapy within 10 days of treatment

Jeffrey M. Cochran
David R. Busch
Anaïs Leproux
Zheng Zhang
Thomas D. O'Sullivan
Albert E. Cerussi
Philip M. Carpenter
Rita S. Mehta
Darren Roblyer
Wei Yang
Keith D. Paulsen
Brian Pogue

Shudong Jiang
Peter A. Kaufman
So Hyun Chung
Mitchell Schnall
Bradley S. Snyder
Nola Hylton
Stefan A. Carp
Steven J. Isakoff
David Mankoff
Bruce J. Tromberg
Arjun G. Yodh

SPIE.

Tissue oxygen saturation predicts response to breast cancer neoadjuvant chemotherapy within 10 days of treatment

Jeffrey M. Cochran,[a,*] David R. Busch,[b] Anaïs Leproux,[c] Zheng Zhang,[d] Thomas D. O'Sullivan,[c] Albert E. Cerussi,[c] Philip M. Carpenter,[e] Rita S. Mehta,[f] Darren Roblyer,[g] Wei Yang,[h] Keith D. Paulsen,[i] Brian Pogue,[i] Shudong Jiang,[i] Peter A. Kaufman,[j] So Hyun Chung,[a] Mitchell Schnall,[k] Bradley S. Snyder,[l] Nola Hylton,[m] Stefan A. Carp,[n] Steven J. Isakoff,[o] David Mankoff,[p] Bruce J. Tromberg,[c] and Arjun G. Yodh[a]

[a]University of Pennsylvania, Department of Physics and Astronomy, Philadelphia, Pennsylvania, United States
[b]University of Texas Southwestern, Department of Anesthesiology and Pain Management, Dallas, Texas, United States
[c]University of California, Beckman Laser Institute and Medical Clinic, Irvine, California, United States
[d]Brown University School of Public Health, Department of Biostatistics and Center for Statistical Sciences, Providence, Rhode Island, United States
[e]University of Southern California, Keck School of Medicine, Department of Pathology, Los Angeles, California, United States
[f]University of California Irvine, Department of Medicine, Irvine, California, United States
[g]Boston University, Department of Biomedical Engineering, Boston, Massachusetts, United States
[h]University of Texas MD Anderson Cancer Center, Department of Diagnostic Radiology, Houston, Texas, United States
[i]Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire, United States
[j]Dartmouth-Hitchcock Medical Center, Department of Hematology and Oncology, Lebanon, New Hampshire, United States
[k]University of Pennsylvania, Department of Radiology, Philadelphia, Pennsylvania, United States
[l]Brown University School of Public Health, Center for Statistical Sciences, Providence, Rhode Island, United States
[m]University of California, Department of Radiology, San Francisco, California, United States
[n]Massachusetts General Hospital, Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Boston, Massachusetts, United States
[o]Massachusetts General Hospital, Department of Hematology and Oncology, Boston, Massachusetts, United States
[p]University of Pennsylvania, Division of Nuclear Medicine, Department of Radiology, Philadelphia, Pennsylvania, United States

Abstract. Ideally, neoadjuvant chemotherapy (NAC) assessment should predict pathologic complete response (pCR), a surrogate clinical endpoint for 5-year survival, as early as possible during typical 3- to 6-month breast cancer treatments. We introduce and demonstrate an approach for predicting pCR within 10 days of initiating NAC. The method uses a bedside diffuse optical spectroscopic imaging (DOSI) technology and logistic regression modeling. Tumor and normal tissue physiological properties were measured longitudinally throughout the course of NAC in 33 patients enrolled in the American College of Radiology Imaging Network multicenter breast cancer DOSI trial (ACRIN-6691). An image analysis scheme, employing z-score normalization to healthy tissue, produced models with robust predictions. Notably, logistic regression based on z-score normalization using only tissue oxygen saturation (S_tO_2) measured within 10 days of the initial therapy dose was found to be a significant predictor of pCR (AUC = 0.92; 95% CI: 0.82 to 1). This observation suggests that patients who show rapid convergence of tumor tissue S_tO_2 to surrounding tissue S_tO_2 are more likely to achieve pCR. This early predictor of pCR occurs prior to reductions in tumor size and could enable dynamic feedback for optimization of chemotherapy strategies in breast cancer. © *The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JBO.24.2.021202]

Keywords: biomedical optics; neoadjuvant chemotherapy; diffuse optical spectroscopy; breast cancer; therapy monitoring; translational imaging.

Paper 180342SSR received Jun. 8, 2018; accepted for publication Aug. 30, 2018; published online Oct. 19, 2018.

1 Introduction

Neoadjuvant chemotherapy (NAC) is a widely used treatment method for breast cancer that permits increased conservation of breast tissue during tumor resection and limits the need for axillary node treatment and surgery.[1] In addition, pathologic complete response (pCR) to NAC, defined as no residual invasive carcinoma, has been correlated with improved survival compared to incomplete response.[2,3] Unfortunately, this assessment occurs after the completion of NAC. The ability to predict response to NAC at an earlier timepoint during chemotherapy, by contrast, could enable physicians to dynamically optimize the treatment regimen, thereby avoiding unnecessary therapy doses, reducing tissue damage, and improving patient outcomes.

NAC response is typically evaluated with physical exams and radiologic imaging in current clinical practice. Unfortunately, these methods are inadequate predictors of pCR.[4–6] Magnetic resonance imaging (MRI) provides better correlation with pathology than mammography or ultrasound.[7] Broadly, functional monitoring techniques offer significantly improved correlation with response relative to structural imaging modalities. Magnetic resonance spectroscopy (MRS),[8] contrast-enhanced

MRI,[9] and positron emission tomography (PET),[10–12] have predictive value with respect to pCR, but MRI, MRS, and PET all have practical constraints, which limit the frequency of monitoring in clinical care. These limitations include cost, the use of contrast agents, and ionizing radiation for PET.

The present contribution investigates the utility of diffuse optical monitoring for prediction of pCR during NAC and adds an analysis to prior reports of a multicenter trial.[13] Briefly, diffuse optical techniques measure functional hemodynamic properties of tissue with nonionizing near-infrared radiation. These optical methods are relatively low cost and can be employed at the bedside. Furthermore, the technology offers a quantitative tool to predict treatment outcome based on longitudinal measurements during therapy.[14,15] Diffuse optical spectroscopic imaging (DOSI) and tomography (DOT) probe deeply, i.e., several centimeters, into tissue and provide information about tissue optical absorption (μ_a) and reduced optical scattering (μ_s'), from which deoxygenated-(HHb) and oxygenated-hemoglobin (HbO$_2$) concentrations, as well as lipid and water (H$_2$O) concentrations can be deduced.[16,17] The concentrations of HHb and HbO$_2$, in turn, are readily utilized to calculate total hemoglobin concentration (Hb$_T$) and tissue oxygen saturation (S$_t$O$_2$). These parameters have been shown to discriminate malignant from healthy tissue in the breast,[18–23] and several studies have employed DOSI techniques to explore functional changes in malignancies during NAC and have correlated these changes with response to therapy.[13,24–35]

We recently reported the first results of ACRIN-6691, an American College of Radiology Imaging Network (ACRIN) multicenter clinical trial of patients monitored longitudinally by DOSI throughout their NAC regimen.[13] The primary aim of ACRIN-6691 was to evaluate whether a change in a particular DOSI endpoint, the tissue optical index (TOI), could be used to predict a clinical endpoint, pCR, by the midpoint of NAC, ~2 to 3 months after the first infusion. The TOI combines tissue deoxyhemoglobin concentration (HHb), water, and lipid into a single index (see Sec. 2). In that initial study, we reported significant reductions in tumor to normal (T/N) TOI ratios for pCR subjects. A 40% or greater change in this parameter at midpoint, combined with baseline tumor S$_t$O$_2$ greater than median values (77%) was shown to be a promising predictor of pCR (AUC = 0.83; 95% CI: 0.63 to 1).[13]

In this study, we explore the ACRIN-6691 secondary aim of predicting pCR much earlier in the 3- to 6-month NAC cycle by examining DOSI response parameters within 10 days of therapy initiation. To address this goal, we develop and retrospectively apply z-score normalization[21] and a logistic regression algorithm[36] to correlate DOSI-measured parameters of malignant breast lesions to subjects' posttherapy pathologic response status. Our hypothesis is that identification and optimization of this z-score DOSI index could predict pCR to NAC at an early timepoint in the course of therapy, providing significant potential for clinical utility.

2 Materials and Methods

2.1 Trial Design and Subjects

Data for this study were collected during the ACRIN 6691 multisite trial using a DOSI instrument developed at the University of California, Irvine.[13] Subjects provided written informed consent, and the HIPAA-compliant protocol and

Table 1 Subject and tumor characteristics. Demographic, histological, and immunohistochemical data are provided for all subjects and divided into complete responder (pCR) and noncomplete responder (non-pCR) groups. For histological information, IDC refers to invasive ductal carcinoma, ILC refers to invasive lobular carcinoma, and DCIS is ductal carcinoma in-situ. ER, PR, and Her2 represent estrogen receptor, progesterone receptor, and human epidermal growth factor receptor status, respectively.

	pCR ($N = 15$)	Non-pCR ($N = 18$)
Age (years)		
Mean $\pm$ SD (range)	49.0 $\pm$ 11.6 (30 to 67)	49.4 $\pm$ 10.9 (28 to 66)
Menopausal status, *n* (%)		
Pre	5 (33)	9 (50)
Peri	1 (7)	2 (11)
Post	9 (60)	7 (39)
Maximum tumor size (mm)		
Mean $\pm$ SD (range)	37.9 $\pm$ 22.8 (12 to 95)	37.5 $\pm$ 18.1 (11 to 75)
Histological findings, *n* (%)		
IDC	9 (60)	12 (67)
ILC	0 (0)	1 (6)
IDC/DCIS	4 (27)	5 (28)
IDC/ILC	1 (7)	0 (0)
Unknown	1 (7)	0 (0)
ER status, *n* (%)		
Positive	5 (33)	16 (89)
Negative	8 (53)	2 (11)
Unknown	2 (13)	0 (0)
PR status, *n* (%)		
Positive	5 (33)	11 (61)
Negative	8 (53)	7 (39)
Unknown	2 (13)	0 (0)
Her2 status, *n* (%)		
1	5 (33)	4 (22)
2	1 (7)	9 (50)
3	5 (33)	1 (6)
Unknown	4 (27)	4 (22)
Molecular subtype, *n* (%)		
Her2 positive	3 (20)	1 (6)
HR positive	0 (0)	2 (11)
Luminal A	0 (0)	3 (17)
Luminal B	6 (40)	11 (61)
Triple negative	4 (27)	1 (6)
Unknown	2 (13)	0 (0)

informed consent were approved by the American College of Radiology Institutional Review Board, the NCI Cancer Therapy Evaluation Program, and each site's Institutional Review Board. All 60 enrolled subjects were females between the ages of 28 and 67 with biopsy-confirmed invasive ductal carcinomas and/or invasive lobular carcinomas of at least 2 cm in length along the greatest dimension. For each subject, the chemotherapy regimen was determined by the subject's physician. Chemotherapy type was not controlled in this study, except that regimens were required to include at least one cytotoxic chemotherapeutic agent.

pCR to therapy was defined as no residual invasive primary carcinoma without regard to residual lymph node disease and was determined for each subject from postsurgery pathology reports. Subjects that achieved partial response were not distinguished from nonresponders because of statistical considerations, i.e., sample size, and due to the previously

reported correlation between complete response and improved survival.[2,3] Table 1 contains demographic information, as well as tumor histology and immunohistochemistry for complete and noncomplete responders.

A number of enrolled subjects were excluded from the final data set. Of these, three subjects withdrew from the study. An additional 13 subjects were not included in the imaging analysis because of the following DOSI scan issues: mandatory baseline DOSI was not performed ($n = 1$), baseline DOSI was nonevaluable ($n = 8$), mandatory midtherapy DOSI was not performed ($n = 3$), or too few normal region points were available ($n = 1$). A DOSI scan was considered nonevaluable in case of unrealistic physiological values or incorrect instrument configuration. This decision was made on blinded, deidentified data using instrument calibration and raw data QC reports.[13] A flowchart for this exclusion process can be found in Fig. 6 in the Appendix.
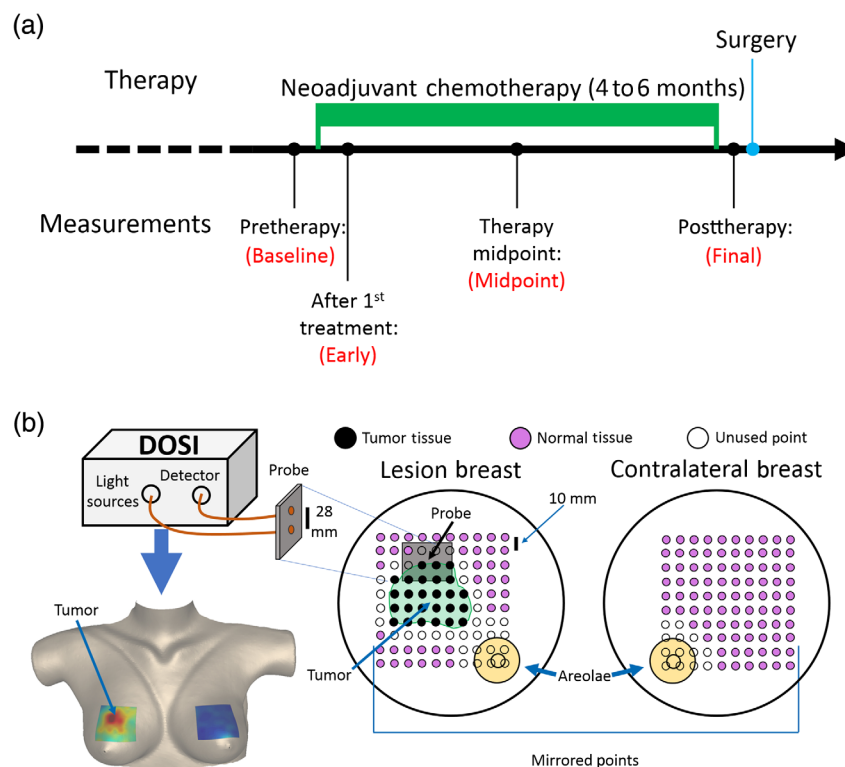

Fig. 1 Timeline and schematic of DOSI measurement during NAC. (a) Each enrolled subject underwent NAC for a period of 4 to 6 months. DOSI measurements were made at four timepoints throughout the course of therapy: (1) baseline—prior to the administration of therapy, (2) early—5 to 10 days after the first dose of therapy, (3) midpoint—the midpoint of the therapy regimen, (4) final—at least 7 days after the final dose of therapy and prior to tumor resection. Note that some subjects are missing data at one or more of the nonbaseline timepoints, and the measurements at the final timepoint were not used due to their limited predictive utility. (b) Top left: DOSI instrument and probe. Right: a grid of points, over a surface area ranging from 7 cm × 7 cm to 15 cm × 16 cm, were measured on the lesion-bearing breast. This grid was chosen to encompass both the tumor and a portion of the surrounding healthy tissue. The grid of points was marked using a transparency, which was then used to mirror the grid for measurements on the contralateral breast. The transparency was also used to ensure consistent measurement locations across all timepoints. The tumor region was chosen to be all contiguous points with magnitude greater than half of the maximum TOI measurement. The tumor-bearing breast normal region was defined as all points outside the tumor region and areola, excluding a 1-cm margin around both the tumor and areola. The contralateral breast normal region was defined as all measured points, excluding the areola and a 1-cm margin around the areola. Bottom left: a sample DOSI image of the TOI contrast mapped onto a 3-D breast surface (see Sec. 2).

2.2 Optical Imaging Methods

The DOSI instrument used in this study combines multispectral frequency-domain and broadband diffuse optical spectroscopy to measure tissue concentrations of oxygenated hemoglobin (HbO_2), deoxygenated hemoglobin (HHb), water (H_2O), and lipid, as well as the tissue scattering amplitude (A) and power (b), as defined by a simplified Mie scattering model.[37] The combination of these measured parameters permits calculation of total tissue hemoglobin concentration ($Hb_T = HbO_2 + HHb$), tissue oxygen saturation ($S_tO_2 = HbO_2/Hb_T$), and the tissue reduced scattering coefficient (μ'_s). For a full description of the DOSI method and instrument performance in the multicenter ACRIN-6691 trial, see Ref. 38.

The DOSI instrument measured subjects using a handheld probe (handpiece) placed in contact with the patient's breast. Four timepoints were acquired throughout the course of each patient's NAC regimen[13] (see Fig. 1). The first measurement (baseline) occurred prior to the first dose of chemotherapy. The second measurement, which is referred to as the early measurement timepoint, was performed between 5 and 10 days after the first chemotherapy treatment. The third measurement (midpoint) occurred in the middle of the therapy regimen, and a final measurement was made after the completion of therapy but prior to tumor resection. During each subject's baseline measurement, a grid of ∼50 to 240 points that encompassed both the palpated tumor region and the surrounding normal tissue was measured on the lesion-bearing breast. A mirrored grid of points was measured on the contralateral breast (see Fig. 1). These measurement grids were recorded using a hand-marked transparency film that was produced for each subject in order to guide DOSI handpiece placement to the grid points during each measurement session as previously described.[13]

2.3 Statistical and Analytic Methods

In this study, we trained a logistic regression algorithm to discriminate between responders and nonresponders based on DOSI-measured parameters (list of parameters available in Table 2 in the Appendix). The tumor region for each subject was determined using a TOI [TOI = (HHb · H_2O)/lipid]. This TOI parameter has been empirically shown to differentiate malignant tissue from normal tissue in the breast.[19] The full-width-at-half-maximum contour around the point of maximum TOI in the baseline measurement of the lesion-bearing breast was designated as the edge of the tumor region. This region remained constant throughout all longitudinal measurements for a given subject. The normal region on the lesion-bearing breast was defined as all points outside the tumor region excluding the areola and 1-cm margins around the areola and tumor region (see Fig. 1). These margins were not included in the normal region to limit signal contamination due to the partial volume effect.

In practice, significant inter- and intrasubject variation in optically measured physiological parameters of the breast can arise,[21,39] and these systemic variations can bias the logistic regression. Moreover, the optically measured tissue parameters are not normally distributed (see Fig. 2).

To remedy these issues, we introduce and employ a z-score normalization method to define target variables for prediction of pathologic response. Briefly, the natural logarithm of each data point is first taken. Then, the mean and standard deviation of a
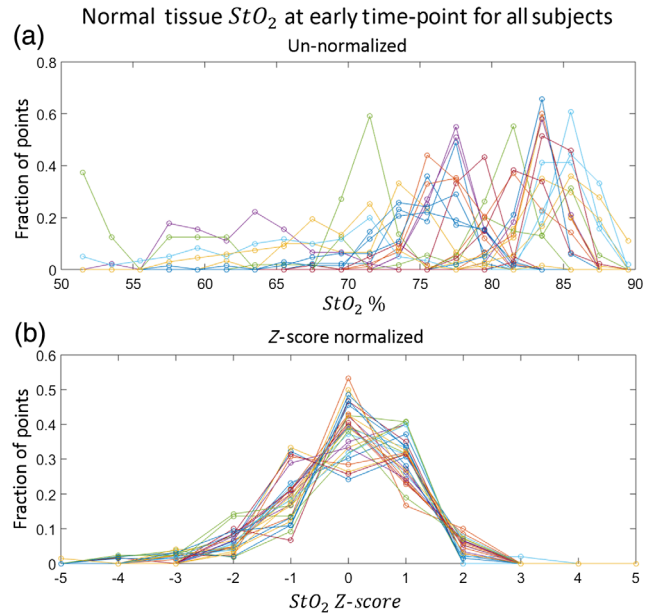

Fig. 2 Histograms of the early timepoint normal tissue S_tO_2 for all subjects. (a) Fractional histograms of the unnormalized S_tO_2 of the normal tissue on the tumor-bearing breast at the early timepoint for each subject. Each line represents a different subject. (b) Fractional histograms of the z-score normalized log-transformed S_tO_2 data of the normal tissue on the tumor-bearing breast at the early timepoint for each subject. Each line represents a different subject. Note that with the z-score normalization, the distributions for all subjects have the same mean and an approximately Gaussian distribution. This effect is consistent across all measured parameters and timepoints.

normal (healthy) region of tissue are used to transform raw tumor data into z-score data as in

$$Z_j = \frac{\ln X_j - \langle \ln X_{j_{Norm}} \rangle}{\sigma[\ln X_{j_{Norm}}]}. \qquad (1)$$

Here, X_j is the unnormalized j'th measured parameter in the tumor region, $X_{j_{Norm}}$ is the unnormalized j'th measured parameter in the normal (healthy) region of the tumor-bearing breast, $\langle \ln X_{j_{Norm}} \rangle$ represents the mean over all points in the normal (healthy) region, and $\sigma[\ln X_{j_{Norm}}]$ represents the standard deviation over all points in the normal (healthy) region. Z_j is then the tumor region z-score relative to the healthy tissue for the j'th parameter. Each Z_j parameter was averaged over all spatial points in the tumor region for a given subject and timepoint. As a result, the logistic regression algorithms can utilize a single tumor quantity for each subject, for each timepoint, and for each measured parameter.

Thus every predictor data point used in the regression model is measured in units of standard deviations from the mean of a given parameter in healthy tissue. In addition to transforming all parameters to be approximately the same magnitude, this method better accounts for the intersubject systemic variations by finding the difference of each parameter from the mean value of the normal (healthy) tissue. It also more fully accounts for intrasubject variation in healthy tissue by normalizing with the healthy tissue standard deviation. A concrete example of this statistical transformation scheme is shown in Fig. 2 for early timepoint tissue oxygen saturation. In this study,

we explored z-score normalization schemes that defined the normal region as either the healthy breast (excluding the areola) or all tissue on the lesion breast outside a certain margin of the tumor region (excluding the areola). See Fig. 1 for a graphical representation of these different normalization regions.

In the logistic regression framework, a response parameter for a given model is defined as

$$\mathcal{R}^i = \beta_o + \sum_{j=1}^{N_j} \beta_j \cdot Z_j^i. \tag{2}$$

Here, $\mathcal{R}^i$ is the given model's log odds of response for the i'th subject, β_o is the intercept term of the fitted weight vector, β_j is the weighting term for the j'th measured parameter used in the model, Z_j^i is the z-score for the j'th measured parameter of the i'th subject, and N_j is the number of parameters used in a particular model. The full weight vector $\vec{\beta}$ is

$$\vec{\beta} = [\beta_o, \beta_1, \ldots, \beta_{N_j}]. \tag{3}$$

The $\vec{\beta}$ weight vector is fit using MATLAB's native logistic regression function, mnrfit.[40] The response parameter $\mathcal{R}$ can then be transformed into a probability of response parameter $\mathcal{P_R}$ using a logistic function

$$\mathcal{P_R} = \frac{1}{1 + e^{-\mathcal{R}}}. \tag{4}$$

The parameter $\mathcal{P_R}$ represents the probability that a subject will achieve pCR. It has a range from 0 to 1, and it can readily be used to predict each subject's status as either a pathologic complete responder, or noncomplete responder, depending on threshold levels.

Because we are working with a small dataset, we employed a leave-one-out validation protocol[41] to test the regression model. Briefly, a series of logistic regression models are created for each parameter set we wish to test, and each of these models leaves one of the subjects out of the dataset (see Fig. 3). The weight vector created by each of these models $\vec{\beta}^i$ is the weight vector created when the i'th subject is left-out; it is used to produce a probability of response prediction for the i'th subject ($P_\mathcal{R}^i$), which is independent of the $\vec{\beta}^i$ model. This well-known
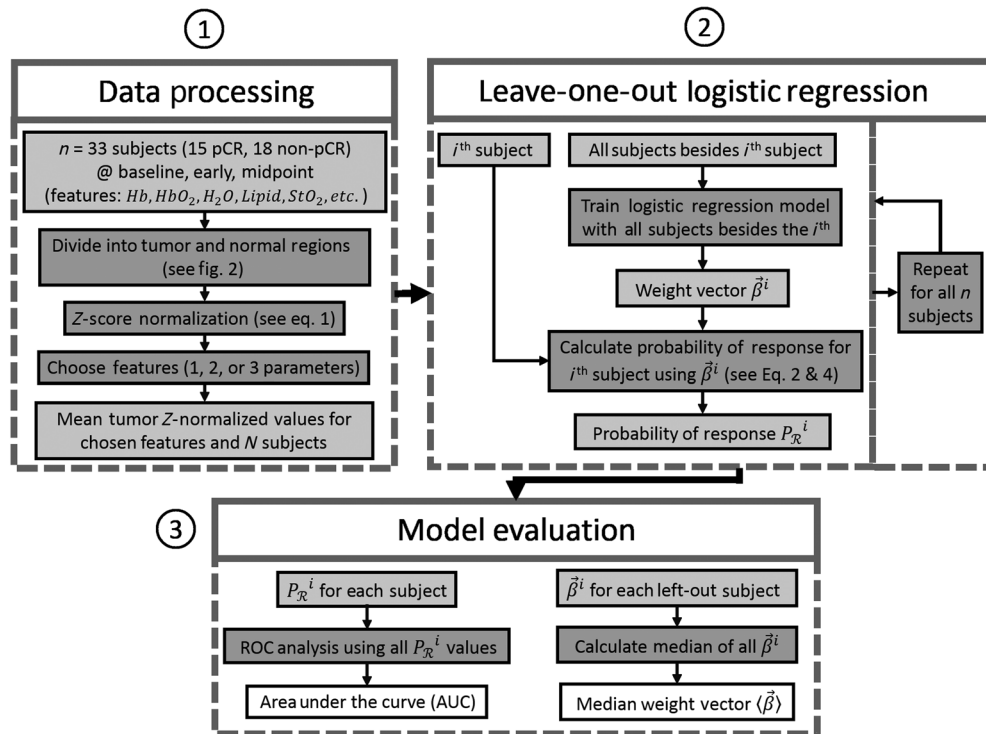

Fig. 3 Data analysis flowchart. (1) Data processing—measured quantities at all spatial points and all n subjects across the first three timepoints are first divided into tumor and normal (healthy) regions (see Fig. 1). All tumor points are then z-score normalized to their respective normal (healthy) regions [see Eq. (1)], and the mean is taken for a given subject and timepoint. Finally, one-, two-, or three-model parameters are chosen from among the combinations of measured quantities and timepoints as model inputs. (2) Leave-one-out logistic regression—a set of n logistic regression algorithms are performed, each of which leaves out a single subject from the training data and produces a $\vec{\beta}^i$ weight vector. Each $\vec{\beta}^i$ is then used to calculate the probability of response for the subject left out of the given training set [see Eqs. (2) and (4)]. (3) Model evaluation—ROC analysis is performed using the calculated $P_\mathcal{R}^i$ values to determine the AUC and a median weight vector $\langle\vec{\beta}\rangle$ is calculated from the n resulting $\vec{\beta}^i$ vectors.

approach provides the most robust and least biased validation given our sample size, which precludes the use of a significantly large independent test set.[41] For completeness, we compared the leave-one-out protocol to other methods. For example, we also tested k-fold cross validation with $k = 3$, 5, and 10; these schemes produced $\vec{\beta}$ vectors and AUC values that were similar to those of the leave-one-out protocol.

The quality of the resultant models was empirically determined using DeLong's method for the area under the curve (AUC) and 95% confidence interval of a receiver-operating-characteristic (ROC) analysis graph for the $\mathcal{P}_\mathcal{R}$ parameter.[42] The ROC analysis is performed using each of the individual leave-one-out models, and the reported weight vector $\langle\vec{\beta}\rangle$ will be the median $\vec{\beta}$ from the series of models created for each parameter set, with the interquartile range (IQR) of these models reported as uncertainty. We also calculated a single-logistic regression model run across the entire dataset; it produced very similar $\vec{\beta}$ vectors to the median $\vec{\beta}$ vector approach described above.

Models based on the z-scores of each single measured parameter (HHb, HbO$_2$, lipid, H$_2$O, Hb$_T$, S$_t$O$_2$, and TOI) at the baseline, early, and midpoint timepoints were tested, and the most predictive models were chosen using the AUC value as a criterion. To explore any additional benefit from multivariate models, combinations of two and three measured parameters were also evaluated. Models with more than three parameters were not tested to avoid overfitting. Data from the final timepoint were not used because our focus in this work is on early diagnosis.

Other, more commonly used, normalization methods were also tested to demonstrate the improvement in predictive ability provided by z-score normalization. These comparisons included tumor-to-normal ratio normalization without information about the normal tissue heterogeneity, as well as raw tumor physiological values without normalization, and baseline-normalized values, which represent changes in the measured parameters over the course of the therapy regimen. All statistical analysis was performed using MATLAB R2015a (The Mathworks, Inc., Natick, Massachusetts, USA).[40]

3 Results

The final data set was derived from $n = 33$ subjects who had complete data sets at the baseline and midpoint timepoints. For models that used measured parameters from the early timepoint, slightly fewer subjects were used ($n = 29$) due to missing data at this timepoint. All subjects had biopsy-confirmed invasive carcinomas and underwent an NAC regimen determined by their physicians.[13]

For the logistic regression algorithm, z-score normalization to the healthy tissue on the lesion breast, as opposed to normalization to the contralateral breast, produced more predictive models. Recall that we derive z-score data for multiple data types (HbO$_2$, HHb, Hb$_T$, S$_t$O$_2$, H$_2$O, and lipid) at multiple timepoints (baseline, early, and midpoint) (all data available in Table 3 in the Appendix). The single best regression model used only the early timepoint tissue oxygen saturation (eS$_t$O$_2$). The weight vector for this model was $\langle\vec{\beta}\rangle = [\beta_o = 0.79 \pm 0.09, \beta_{eS_tO_2} = 2.29 \pm 0.04]$. This finding suggests that, at early timepoints, tumors that are not hypoxic relative to the surrounding normal tissue, or tumors that are only slightly hypoxic and within the normal region's confidence interval, are more likely

to be pathologic complete responders to NAC. By contrast, tumors that were significantly hypoxic relative to the normal tissue were likely to be nonresponders (see Fig. 4 for data summary in traditional units). When ROC analysis was performed, this model produced an AUC = 0.92 with a 95% confidence interval of AUC = 0.82 to 1 (see Fig. 5). Additionally, the small uncertainties of the $\langle\vec{\beta}\rangle$ components, relative to the median, indicate that the fitted $\langle\vec{\beta}\rangle$ did not vary significantly across the leave-one-out validation protocol.

Two- and three-parameter models did not improve upon the single-parameter model AUC. Higher-order models, e.g., four-parameter, were not considered in order to avoid overfitting of the data.

Notably, in addition to the early timepoint oxygen saturation, a two-parameter model using only baseline data provided an AUC = 0.83 with a 95% confidence interval of AUC = 0.70 to 0.97, thus enabling an even earlier prediction of a subject's pCR status, albeit with lower accuracy than the early timepoint oxygen saturation. This two-parameter model incorporated the baseline oxygen saturation (bS$_t$O$_2$) and water concentration (bH$_2$O), and the median weight vector was $\langle\vec{\beta}\rangle = [\beta_o = 0.14 \pm 0.09, \beta_{bS_tO_2} = 1.69 \pm 0.06, \beta_{bH_2O} = 0.65 \pm 0.03]$. Again, the uncertainties in the $\langle\vec{\beta}\rangle$ components for H$_2$O and S$_t$O$_2$ are small, signifying a consistent fitted model across
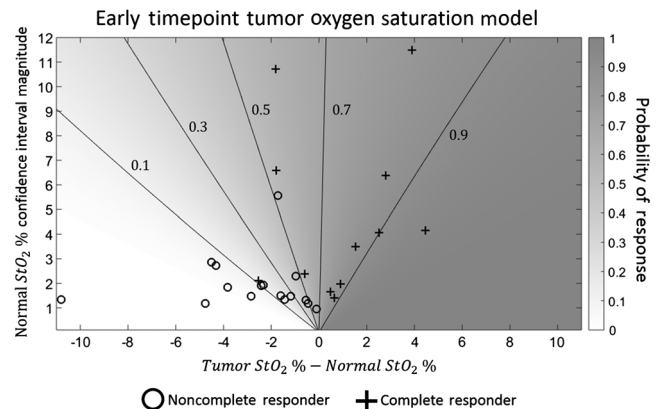

Fig. 4 Tumor and normal S$_t$O$_2$ versus probability of response. This graph shows the probability of response predicted by the regression model using only early timepoint S$_t$O$_2$ (see Fig. 5). Contour lines of constant probability are also included. The probability of response (shading) is plotted versus the difference between the absolute tumor region percent oxygen saturation and the absolute normal region percent oxygen saturation (horizontal axis), and the size of the confidence interval for the absolute normal region oxygen saturation, corresponding to one standard deviation in the log-transformed data (vertical axis). Note that the oxygen saturation in this figure is not log-transformed or z-score normalized. Each cross represents a subject that was a pathologic complete responder while each circle indicates a nonresponding subject. All subjects that had tumor regions with absolute oxygen saturations that were higher than their normal regions achieved pCR. Subjects whose tumor regions were only slightly hypoxic relative to their normal regions were more likely to achieve pCR if the subjects' normal regions had larger confidence intervals. These observations indicate that a subject is likely to be a pathologic complete responder if the oxygen saturation of the tumor region is either higher than that of the normal region or well within the normal region's confidence interval. A subject whose tumor was significantly hypoxic relative to the normal tissue was likely to be a nonresponder.
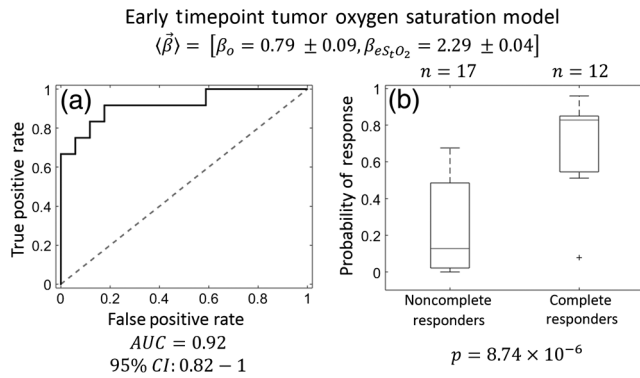
Fig. 5 Early timepoint oxygen saturation prediction model. The model providing the best predictions used the early timepoint tissue oxygen saturation (eS_tO_2). The median weight vector $\langle\vec{\beta}\rangle = [\beta_o = 0.79 \pm 0.09, \beta_{eS_tO_2} = 2.29 \pm 0.04]$ indicates that tumors that are not hypoxic relative to the normal tissue on the tumor breast are more likely to be pathologic complete responders to chemotherapy. (a) ROC analysis of eS_tO_2 model—this model produced an AUC = 0.92 (95% CI: 0.82 to 1), indicating excellent predictive value. (b) Boxplots of probability of response—the probability of response boxplots, divided into subjects that achieved pCR ($n = 12$) and subjects that did not achieve pCR ($n = 17$), indicate clear separation between the two groups using this model ($p = 8.74 \times 10^{-6}$ using a two-sided student's t-test). The hinges of the boxplots represent the first and third quartiles of the data, the whiskers represent the range of measurements within a distance 1.5× the IQR, and the cross represents an outlier. Note that there is no overlap between the IQRs of the probability of response of the complete responders and noncomplete responders.

the leave-one-out validation procedure. The fact that $\beta_{bStO_2} > \beta_{bH_2O}$ indicates that the oxygen saturation is a more significant predictor of pCR than water concentration at the baseline timepoint. As with the early timepoint model, subjects with hypoxic tumors were less likely to achieve pCR.

For comparison, additional models were produced that (1) used the contralateral breast for z-score normalization, (2) used tumor-to-normal ratio normalization, i.e., with no information about the standard deviation of the normal region, and (3) used no normalization. With contralateral z-score normalization, instead of z-score normalization to the healthy tissue on the tumor-bearing breast, the aforementioned early (eS_tO_2) and baseline (bS_tO_2 and bH_2O) models had AUC values of 0.67 and 0.64, respectively. With simple tumor-to-normal ratio normalization, the same two-parameter sets produced AUC values of 0.80 and 0.67, and when completely unnormalized data were used, the eS_tO_2 model produced an AUC = 0.75 while the bS_tO_2 and bH_2O model produced an AUC = 0.68. Thus, for these parameter sets, z-score normalization to the healthy tissue in the tumor-bearing breast provided the best results.

4 Discussion

By application of a logistic regression model using z-score normalized DOSI measurements, we derived a robust predictor of response (AUC = 0.92; 95% CI: 0.82 to 1) within the first 10 days after a subject's initial chemotherapy dose. Using an optimally chosen cutoff value of $\mathcal{P}_R = 0.50$, which maximizes the sum of the sensitivity and specificity, this model provided an overall classification accuracy of 86% (25 of 29 subjects), including a positive predictive value of 79% for subjects

predicted to achieve pCR (11 of 14), and a negative predictive value of 93% for subjects predicted to not achieve pCR (14 of 15). Prediction of response at this therapy timepoint was a secondary aim of the ACRIN 6691 protocol[13] and could, with further validation, enable clinicians to modify the patient's therapeutic plan after a single dose. This ability holds potential to improve patient outcomes and prevent unnecessary side effects from ineffective treatments.

The best model indicated that low S_tO_2 at the early timepoint relative to the surrounding normal tissue was predictive of nonresponse to chemotherapy. This observation suggests that tumors that are well-perfused in the early stages of treatment, and therefore are not hypoxic relative to healthy tissue, may receive chemotherapy more efficiently. Such tumors are also typically more responsive to therapy than hypoxic tumors, which often exhibit resistance to treatment.[43,44] Additionally, the lack of hypoxia in complete responders could indicate a decreased oxygen demand due to suppression of tumor metabolism, a condition previously shown to be correlated with response to therapy.[45]

Additionally, the two-parameter model using only the baseline S_tO_2 and water concentration (AUC = 0.83; 95% CI: 0.70 to 0.97) also indicated that higher S_tO_2 is correlated with pCR. Though the AUC value is lower for this model compared to the early timepoint S_tO_2 model, prediction of response prior to the initiation of therapy offers additional clinical utility. These models are also consistent with previous studies, which have observed correlation between pCR and optically measured tissue oxygen saturation prior to the start of therapy[31] and after the first dose.[29]

Previous diffuse optical studies of response to breast cancer NAC have correlated temporal changes in measured physiological parameters with response to treatment.[24–30,32] We compared our technique to this approach in the current study. However, even the most predictive of the models derived in this analysis that used the change in DOSI physiological parameters between the baseline and early timepoints only produced an AUC = 0.63. The temporal change models of S_tO_2, in particular, could be limited by the large intersubject dispersion of the baseline oxygen saturation; this large dispersion prevents the change in S_tO_2 from the baseline to early timepoint from accurately reflecting the oxygenation state of the tumor relative to the normal region. By contrast, the model we have presented in this contribution does not depend on the baseline S_tO_2 and, as such, is not affected by intersubject baseline variation.

Z-score normalization was implemented to place all parameters on the same magnitude scale, which mitigates systemic physiological differences among the subject population and accounts for the systemic effects of chemotherapy. For comparison, we also investigated models that used fully unnormalized data and tumor-to-normal ratio normalization. However, since neither model incorporates healthy tissue standard deviation, neither model accounts for the heterogeneity of normal tissue. With tumor-to-normal ratio normalization, a one-parameter model with early timepoint S_tO_2 produced an AUC = 0.80, and the two-parameter model with baseline timepoint S_tO_2 and H_2O produced an AUC = 0.67. The AUC values for the same models but with no normalization were even lower (AUC = 0.75 and AUC = 0.67, respectively). Thus z-score normalization improves the predictive power of the tissue oxygen saturation logistic regression models.

For completeness, several other models were explored that did not incorporate the baseline or early S_tO_2. Some of these produced predictions of response to therapy that were significant (AUC ≈ 0.75 to 0.80). However, in addition to having lower AUC values, these other models relied on data from the midpoint timepoint, which increases the time-to-prediction of response by ∼2 months. Furthermore, the early timepoint measurements typically occur before significant anatomic changes in tumor size arise. This feature enables the DOSI measurement to sample known tumor tissue more easily; at the midpoint of therapy, by contrast, the tumor size has decreased and signal contamination between the malignant and healthy tissue can occur and limit the ability of DOSI to determine tumor physiological parameters accurately. Note also that the physiological predictions of these other models were consistent with our two primary prediction models.

Another interesting and potentially important finding of the present work is that the best models used z-score normalization to the normal tissue on the lesion breast rather than the contralateral breast. If, instead, the contralateral breast was used, our one-parameter early S_tO_2 model produced an AUC = 0.67, and the two-parameter model with baseline S_tO_2 and H_2O produced an AUC = 0.64. The comparatively better quality of the tumor breast z-score normalized models suggests that measurement of the contralateral breast is less important for early prediction of response to therapy than previously thought. If this is true, then the paradigm could eliminate the need for contralateral measurement and reduce imaging time.

The results we have presented provide evidence for early prediction of response with AUC results that are comparable to other modalities, such as MRI,[46,47] FDG-PET,[11,47,48] and biomarker analysis.[49,50] Some of these studies produced predictions prior to or within the first 10 days of treatment initiation,[48–50] whereas other approaches relied on imaging that occurred either after 6 weeks of NAC,[46] at the midpoint of therapy,[11] or after the completion of NAC.[47] The potential advantage of the logistic regression DOSI model is premised on its unique combination of accurate prediction at an early timepoint in therapy and its portability, low cost, and lack of ionizing radiation.

The primary limitations of this study are the relatively small number of subjects and the highly variable chemotherapy regimens across the subject population. Additionally, the initial study had a fairly high dropout rate,[13] introducing a potential bias into the statistical analysis. The dropout rate is likely to be artificially elevated in this study due to the difficulties inherent in translating an experimental imaging technique into a multisite setting for the first time.[13] We do not anticipate that these issues will affect the DOSI technique moving forward. Finally, although the initial ACRIN 6691 trial was a prospective study, this z-score parameter imaging metric was retrospectively optimized using a standard leave-one-out protocol for multiple potential models. The leave-one-out technique limits overfitting and enhances the generalizability of the prediction metric;[41] it has been extensively used by the cancer community.[31,46,51–53] Of course, a fully prospective validation of this single prediction model, as opposed to the series of models tested here, will be necessary prior to clinical adoption.

Per the first limitation noted above, application of this model to a prospective study with a larger subject population is a natural course of action. Importantly, because the DOSI instrumentation has been shown to provide consistent performance over time, across multiple instruments, and across multiple measurement sites,[38] we anticipate that the weight vector derived for the early timepoint S_tO_2 (see Fig. 5) could be used with z-score normalized measurements in future DOSI studies to calculate a probability of response, i.e., without creating a logistic regression model for each population. In this case, the future study would serve as a direct, independent test set for the results obtained by our current model. Additionally, a logistic regression could also be run on this larger data set to derive an improved prediction model based on a larger training set. If a future study was performed on a significantly different patient population, e.g., patients with tumors in nonbreast tissue, then deriving a weight vector via logistic regression would likely be beneficial.

In addition to providing evidence to further corroborate the results of this pilot investigation, the larger subject population may enable stratification of the subject population by tumor subtype and/or chemotherapy regimen. Our current results are reported for a diverse patient population, various tumor molecular subtypes, and an assortment of chemotherapy regimens (see Table 1). Tumor subtypes may have different levels of tissue oxygen saturation and may respond to chemotherapy differently.[3,54] The physiological mechanisms of chemotherapy regimens also vary. Thus, especially for parameters at the early timepoint, response prediction might be improved by creating individual models for different classes of chemotherapy and/or different tumor subtypes. Also, independent hypoxia biomarkers, such as carbonic anhydrase IX, and measurements of vascular density, such as CD31 staining or DCE-MRI, can be collected at similar timepoints and may enable better understanding of the mechanisms responsible for correlations between tissue oxygen saturation and response. Exploration of these questions should be possible in a larger study.

5 Conclusion

Logistic regression modeling of z-score normalized physiological parameters measured by DOSI was presented and found to predict pCR to NAC. The best model successfully predicted pCR (AUC = 0.92; 95% CI: 0.82 to 1) using tumor and normal tissue oxygen saturation measured within the first 10 days after the initial dose of therapy based on data from the ACRIN 6691 clinical trial.[13] This model suggests that if tumors are hypoxic relative to the surrounding normal tissue, then they are less likely to achieve pCR. These early predictions of therapeutic efficacy are based on quantitative DOSI measurements of tumor (and normal) tissue functional parameters, rather than changes in tumor size, and the z-score normalization of the tumor physiological data yielded improved prediction models compared to tumor-to-normal ratio or unnormalized data. Prospective validation is still needed to confirm these promising results. With this validation, DOSI and logistic regression methods could be used early in NAC to optimize treatment outcomes for individual patients.

Appendix

This appendix contains details about the subject exclusion criteria (Fig. 6) and a more complete accounting of measured z-score parameters for responders and nonresponders at all timepoints (Tables 2 and 3).
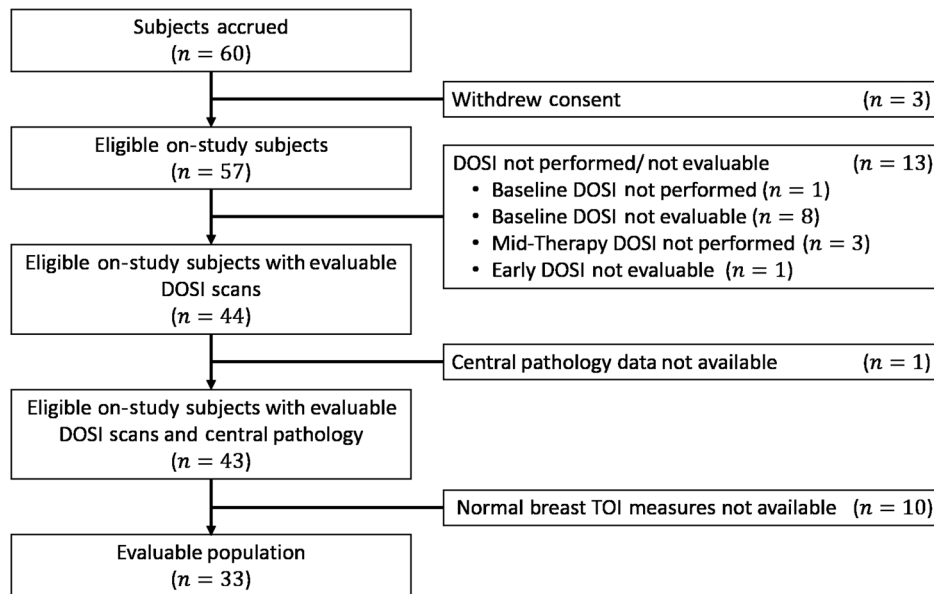
Fig. 6 Subject exclusion chart. Of the 60 subjects accrued for this study, $n = 3$ withdrew consent, $n = 1$ did not have central pathology data, and $n = 10$ were excluded for lack of normal tissue measurement. The other $n = 13$ subjects were excluded due to lack of baseline DOSI measurement ($n = 1$), baseline DOSI measurements that were not evaluable ($n = 8$), lack of midpoint DOSI measurement ($n = 3$), or too few normal region points were available ($n = 1$). This subject population is identical to the population used in the initial ACRIN 6691 study[13] except that one fewer subject was used. This additional excluded subject did have a normal tissue measurement but not a sufficient number of spatial points in the normal region to perform the necessary standard deviation calculation [see Fig. 2 and Eq. (1)].

Table 2 Definitions of measured DOSI parameters and their methods of calculation. HHb, HbO$_2$, lipid, and H$_2$O concentration are all fit directly using the measured intensities throughout the wavelength range. Hb$_T$, StO$_2$, and TOI are all derived from the fit parameters.

Parameter	Meaning	Calculation Method
HHb	Deoxy-hemoglobin concentration	Multispectral fit of absorption
HbO$_2$	Oxy-hemoglobin concentration	Multispectral fit of absorption
Lipid	Lipid concentration	Multispectral fit of absorption
H$_2$O	Water concentration	Multispectral fit of absorption
Hb$_T$	Total hemoglobin concentration	Hb$_T$ = HHb + HbO$_2$
S$_t$O$_2$	Tissue oxygen saturation	S$_t$O$_2$ = HbO$_2$/Hb$_T$
TOI	Tissue optical index	TOI = (HHb · H$_2$O)/lipid

Disclosures

B. J. Tromberg reports intellectual property related to DOSI breast imaging assigned to UC Regents and licensed by Infit, Inc. D. R. Busch and A. G. Yodh have two pending patent applications and A. G. Yodh has two other patents relevant to this work (United States patents 8,082,015 and 6,076,010).

Some of these patents have been licensed to private companies, but neither Dr. Yodh and Dr. Tromberg nor any of the other authors have a financial interest in common with these entities. The other authors have indicated no financial conflicts of interest.

Table 3 Median z-score values with IQRs for each measured parameter and timepoint, separated by pCR status.

	Baseline median z-score (IQR)		Early median z-score (IQR)		Midpoint median z-score (IQR)	
	pCR	Non-pCR	pcR	Non-pCR	pcR	Non-pCR
HbO_2	1.5 (1.1, 1.8)	1.0 (0.4, 1.9)	1.4 (0.8, 1.7)	0.6 (0.1, 1.1)	0.6 (0.4, 1.1)	0.5 (0.2, 1.3)
HHb	2.2 (1.5, 2.7)	1.9 (1.5, 4.0)	1.5 (1.0, 1.9)	1.7 (1.2, 3.4)	1.2 (0.9, 1.6)	1.4 (0.9, 3.0)
Hb_T	1.8 (1.5, 2.4)	1.4 (0.8, 2.2)	1.4 (0.8, 1.8)	0.9 (0.5, 1.3)	1.0 (0.7, 1.4)	0.9 (0.5, 1.6)
S_tO_2	−0.2 (−0.8, 0.1)	−1.8 (−2.3, −0.8)	0.4 (−0.3, 0.5)	−1.2 (−1.7, −0.4)	−0.4 (−1.0, 0.2)	−0.9 (−1.8, −0.6)
H_2O	2.1 (1.5, 2.8)	1.8 (1.2, 3.1)	1.5 (1.3, 2.3)	1.2 (0.8, 3.2)	0.9 (0.4, 2.0)	1.3 (0.9, 2.3)
Lipid	−1.8 (−2.9, −1.4)	−1.4 (−2.5, −0.1)	−1.0 (−1.8, −0.6)	−0.9 (−2.2, −0.1)	−0.7 (−1.1, −0.1)	−0.9 (−1.6, −0.6)
TOI	3.0 (1.7, 3.8)	2.4 (1.6, 4.1)	1.7 (1.2, 2.4)	1.5 (1.2, 4.0)	1.5 (0.9, 2.0)	1.5 (1.2, 3.2)

References

1. G. W. Sledge et al., "Past, present, and future challenges in breast cancer treatment," *J. Clin. Oncol.* **32**(19), 1979–1986 (2014).
2. P. Rastogi et al., "Preoperative chemotherapy: updates of National Surgical Adjuvant Breast and Bowel Project Protocols B-18 and B-27," *J. Clin. Oncol.* **26**(5), 778–785 (2008).
3. L. J. Esserman et al., "Pathologic complete response predicts recurrence-free survival more effectively by cancer subset: results from the I-SPY 1 TRIAL—CALGB 150007/150012, ACRIN 6657," *J. Clin. Oncol.* **30**(26), 3242–3249 (2012).
4. L. D. Feldman et al., "Pathological assessment of response to induction chemotherapy in breast cancer," *Cancer Res.* **46**(5), 2578–2581 (1986).
5. M. A. Helvie et al., "Locally advanced breast carcinoma: accuracy of mammography versus clinical examination in the prediction of residual disease after chemotherapy," *Radiology* **198**(2), 327–332 (1996).
6. S. J. Vinnicombe et al., "Primary breast cancer: mammographic changes after neoadjuvant chemotherapy, with pathologic correlation," *Radiology* **198**(2), 333–340 (1996).
7. E. Yeh et al., "Prospective comparison of mammography, sonography, and MRI in patients undergoing neoadjuvant chemotherapy for palpable breast cancer," *AJR Am. J. Roentgenol.* **184**(3), 868–877 (2005).
8. S. Meisamy et al., "Neoadjuvant chemotherapy of locally advanced breast cancer: predicting response with in vivo (1)H MR spectroscopy—a pilot study at 4 T," *Radiology* **233**(2), 424–431 (2004).
9. N. M. Hylton et al., "Locally advanced breast cancer: MR imaging for prediction of response to neoadjuvant chemotherapy—results from ACRIN 6657/I-SPY TRIAL," *Radiology* **263**(3), 663–672 (2012).
10. D. A. Mankoff et al., "Changes in blood flow and metabolism in locally advanced breast cancer treated with neoadjuvant chemotherapy," *J. Nucl. Med.* **44**(11), 1806–1814 (2003).
11. G. M. McDermott et al., "Monitoring primary breast cancer throughout chemotherapy using FDG-PET," *Breast Cancer Res. Treat.* **102**(1), 75–84 (2007).
12. O. Humbert et al., "Identification of biomarkers including 18FDG-PET/CT for early prediction of response to neoadjuvant chemotherapy in triple-negative breast cancer," *Clin. Cancer Res.* **21**(24), 5460–5468 (2015).
13. B. J. Tromberg et al., "Predicting responses to neoadjuvant chemotherapy in breast cancer: ACRIN 6691 trial of diffuse optical spectroscopic imaging (DOSI)," *Cancer Res.* **76**(20), 5933–5944 (2016).
14. B. J. Tromberg et al., "Assessing the future of diffuse optical imaging technologies for breast cancer management," *Med. Phys.* **35**(6), 2443–2451 (2008).
15. T. D. O'Sullivan et al., "Diffuse optical imaging using spatially and temporally modulated light," *J. Biomed. Opt.* **17**(7), 071311 (2012).
16. D. R. Leff et al., "Diffuse optical imaging of the healthy and diseased breast: a systematic review," *Breast Cancer Res. Treat.* **108**(1), 9–22 (2008).
17. T. Durduran et al., "Diffuse optics for tissue monitoring and tomography," *Rep. Prog. Phys.* **73**(7), 076701 (2010).
18. L. Spinelli et al., "Characterization of female breast lesions from multi-wavelength time-resolved optical mammography," *Phys. Med. Biol.* **50**(11), 2489–2502 (2005).
19. A. Cerussi et al., "In vivo absorption, scattering, and physiologic properties of 58 malignant breast tumors determined by broadband diffuse optical spectroscopy," *J. Biomed. Opt.* **11**(4), 044005 (2006).
20. R. Choe et al., "Differentiation of benign and malignant breast tumors by in-vivo three-dimensional parallel-plate diffuse optical tomography," *J. Biomed. Opt.* **14**(2), 024020 (2009).
21. D. R. Busch et al., "Computer aided automatic detection of malignant lesions in diffuse optical mammography," *Med. Phys.* **37**(4), 1840–1849 (2010).
22. Q. Fang et al., "Combined optical and x-ray tomosynthesis breast imaging," *Radiology* **258**(1), 89–97 (2011).
23. M. A. Mastanduno et al., "Sensitivity of MRI-guided near-infrared spectroscopy clinical breast exam data and its impact on diagnostic performance," *Biomed. Opt. Express* **5**(9), 3103–3115 (2014).
24. B. J. Tromberg et al., "Imaging in breast cancer: diffuse optics in breast cancer: detecting tumors in pre-menopausal women and monitoring neoadjuvant chemotherapy," *Breast Cancer Res.* **7**(6), 279–285 (2005).
25. A. Cerussi et al., "Predicting response to breast cancer neoadjuvant chemotherapy using diffuse optical spectroscopy," *Proc. Natl. Acad. Sci. U. S. A.* **104**(10), 4014–4019 (2007).
26. S. Jiang et al., "Evaluation of breast tumor response to neoadjuvant chemotherapy with tomographic diffuse optical spectroscopy: case studies of tumor region-of-interest changes," *Radiology* **252**(2), 551–560 (2009).
27. H. Soliman et al., "Functional imaging using diffuse optical spectroscopy of neoadjuvant chemotherapy response in women with locally advanced breast cancer," *Clin. Cancer Res.* **16**(9), 2605–2614 (2010).
28. A. E. Cerussi et al., "Diffuse optical spectroscopic imaging correlates with final pathological response in breast cancer neoadjuvant chemotherapy," *Philos. Trans. A Math. Phys. Eng. Sci.* **369**(1955), 4512–4530 (2011).
29. D. Roblyer et al., "Optical imaging of breast cancer oxyhemoglobin flare correlates with neoadjuvant chemotherapy response one day after starting treatment," *Proc. Natl. Acad. Sci. U. S. A.* **108**(35), 14626–14631 (2011).
30. O. Falou et al., "Diffuse optical spectroscopy evaluation of treatment response in women with locally advanced breast cancer receiving neoadjuvant chemotherapy," *Transl. Oncol.* **5**(4), 238–246 (2012).
31. S. Ueda et al., "Baseline tumor oxygen saturation correlates with a pathologic complete response in breast cancer patients undergoing neoadjuvant chemotherapy," *Cancer Res.* **72**(17), 4318–4328 (2012).
32. D. R. Busch et al., "Optical malignancy parameters for monitoring progression of breast cancer neoadjuvant chemotherapy," *Biomed. Opt. Express* **4**(1), 105–121 (2013).
33. P. G. Anderson et al., "Optical mammography in patients with breast cancer undergoing neoadjuvant chemotherapy: individual clinical response index," *Acad. Radiol.* **24**(10), 1240–1255 (2017).

34. R. Choe et al., "Diffuse optical tomography of breast cancer during neoadjuvant chemotherapy: a case study with comparison to MRI," *Med. Phys.* **32**(4), 1128–1139 (2005).

35. J. M. Cochran et al., "Longitudinal optical monitoring of blood flow in breast tumors during neoadjuvant chemotherapy," *Phys. Med. Biol.* **62**(12), 4637–4653 (2017).

36. D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, "Applied logistic regression," in *Wiley Series in Probability and Statistics*, Wiley, Hoboken, New Jersey (2013).

37. J. R. Mourant et al., "Predictions and measurements of scattering and absorption over broad wavelength ranges in tissue phantoms," *Appl. Opt.* **36**(4), 949–957 (1997).

38. A. Leproux et al., "Performance assessment of diffuse optical spectroscopic imaging instruments in a 2-year multicenter breast cancer trial," *J. Biomed. Opt.* **22**(12), 121604 (2017).

39. N. Shah et al., "Spatial variations in optical and physiological properties of healthy breast tissue," *J. Biomed. Opt.* **9**(3), 534–540 (2004).

40. *MATLAB 2015a*, The MathWorks, Inc., Natick, Massachusetts (2015).

41. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer Series in Statistics, Springer Science & Business Media, New York (2008).

42. E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," *Biometrics* **44**(3) 837–845 (1988).

43. M. Hockel and P. Vaupel, "Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects," *J. Natl. Cancer Inst.* **93**(4), 266–276 (2001).

44. R. K. Jain, "Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy," *Science* **307**(5706), 58–62 (2005).

45. M. Dowsett et al., "Endocrine therapy, new biologicals, and new study designs for presurgical studies in breast cancer," *J. Natl. Cancer Inst. Monogr.* **2011**(43), 120–123 (2011).

46. M. L. Ah-See et al., "Early changes in functional dynamic magnetic resonance imaging predict for pathologic response to neoadjuvant chemotherapy in primary breast cancer," *Clin. Cancer Res.* **14**(20), 6580–6589 (2008).

47. S. H. Park et al., "Comparison of diffusion-weighted MR imaging and FDG PET/CT to predict pathological complete response to neoadjuvant chemotherapy in patients with breast cancer," *Eur. Radiol.* **22**(1), 18–25 (2012).

48. L. Kostakoglu et al., "A phase II study of 3'-Deoxy-3'-18F-fluorothymidine PET in the assessment of early response of breast cancer to neoadjuvant chemotherapy: results from ACRIN 6688," *J. Nucl. Med.* **56**(11), 1681–1689 (2015).

49. A. M. Gonzalez-Angulo et al., "Functional proteomics can define prognosis and predict pathologic complete response in patients with breast cancer," *Clin. Proteomics* **8**(1), 11 (2011).

50. M. Ayers et al., "Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer," *J. Clin. Oncol.* **22**(12), 2284–2293 (2004).

51. A. A. Jazaeri et al., "Gene expression profiles associated with response to chemotherapy in epithelial ovarian cancers," *Clin. Cancer Res.* **11**(17), 6300–6310 (2005).

52. K. Nie et al., "Rectal cancer: assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI," *Clin. Cancer Res.* **22**(21), 5256–5264 (2016).

53. M. Jermyn et al., "Highly accurate detection of cancer in situ with intraoperative, label-free, multimodal optical spectroscopy," *Cancer Res.* **77**(14), 3942–3950 (2017).

54. R. Rouzier et al., "Breast cancer molecular subtypes respond differently to preoperative chemotherapy," *Clin. Cancer Res.* **11**(16), 5678–5685 (2005).

Jeffrey M. Cochran is a postdoctoral fellow in the Department of Physics and Astronomy, University of Pennsylvania. He received his BS degree in physics from Denison University and his PhD in physics from the University of Pennsylvania. His research is focused on developing noninvasive imaging biomarkers of disease using diffuse optical techniques.

David R. Busch is an assistant professor in the Department of Anesthesiology and Pain Management and Department of Neurology and Neurotherapeutics at the University of Texas Southwestern Medical Center. His work focuses on developing noninvasive and minimally invasive optical tools to monitor deep tissues in health and disease, as well as throughout therapy.

Thomas D. O'Sullivan is an assistant professor in the Department of Electrical Engineering, University of Notre Dame. He received a PhD in electrical engineering from Stanford University in 2011 and then was a US Department of Defense Breast Cancer Research Program postdoctoral fellow at the University of California, Irvine Beckman Laser Institute. While at BLI, he also served as a director of the Diffuse Optical Spectroscopy and Imaging Laboratory.

Rita S. Mehta is a health science clinical professor in the Department of Hematology/Oncology, University of California, Irvine. She led a phase-III trial of anastrozole versus anastrozole/fulvestrant in advanced breast cancer (S0226, NEJM-2012) and served as an expert member for the development of ASCO guidelines for advanced breast cancer. She completed four phase-II investigator-initiated trials including the use of trastuzumab in HER-2-positive and dose-dense therapy in triple-negative and inflammatory breast cancer.

Darren Roblyer is an assistant professor of biomedical engineering at Boston University. He received his BS degree in biomedical engineering from Johns Hopkins University in 2004 and received his PhD in bioengineering from Rice University in 2009. He did his postdoctoral work at the Beckman Laser Institute, University of California, Irvine. He is a senior member of SPIE.

Arjun G. Yodh is the James M. Skinner professor of science in the Department of Physics and Astronomy, University of Pennsylvania. He directs the Laboratory for Research on the Structure of Matter, an interdisciplinary materials institute that hosts the Penn MRSEC. His biomedical research is oriented toward diffuse optical imaging and monitoring, aiming to demonstrate the potential of these tools for functional imaging/monitoring in brain, breast, and muscle.

Biographies for the other authors are not available