Neoadjuvant Chemotherapy Monitoring of Breast Cancer Patients with Diffuse Optical Tomographic Imaging

Jacqueline E. Gunther

Submitted in partial fulfillment of the
Requirements for the degree of Doctor of Philosophy
In the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2016

ABSTRACT

The overall goal of this thesis was to determine whether optical tomographic imaging can be employed to predict treatment outcome in women with breast cancer (BC) who undergo neoadjuvant chemotherapy (NACT). NACT is widely applied as a standard treatment for patients with newly diagnosed operable invasive BC. Only about 13-30% of women have a response to this treatment. Furthermore, NACT is an expensive and toxic treatment that takes several months to completely administer. In order to know the response of the patient, physicians usually need to wait until the months of NACT has finished and the patient has undergo surgery in which they receive the pathology. If the long-term treatment response could be predicted early into the treatment regimen, the patient would be relieved of any unnecessary side effects and alternative treatments could be initiated. We have used a novel dynamic DOT system to study the effects of targeted NACT. Unlike X-ray imaging, which requires potentially harmful ionizing radiation, DOT can be applied without side effects, which is particularly important in the case of multiple imaging sessions to be performed over the course of treatment. We have tracked 40 subjects and imaged them at 6 different time points during their NACT. For this study, two different types of data were collected: static (single 3D image) and dynamic (3D movies). The combination of the data may be used to accurately determine the response of the patient. With non-invasive, non-ionizing DOT imaging we have been able to determine within two weeks if the patient will respond to treatment with an accuracy as high as 94.1%.

Acknowledgments

I need to thank my advisor Dr. Hielscher, who allowed me to join his lab straight out of undergrad and gave me the daunting responsibility of managing the breast cancer project. I never thought I would end up at an Ivy League school to get a degree, let alone my PhD. He has been an amazing mentor and teacher and I am eternally grateful to him for giving me this opportunity.

I also need to thank Dr. Hyun Keol Kim, whose constant guidance and assistance during my PhD has been much appreciated. I cannot thank him enough for patiently listening and helping me with my project and any issues that I have had with research or school.

Thank you so much to the rest of my committee, Dr. Liane, Dr. Konofagou, and Dr. Hershman for their support and advice on this project. I especially need to thank Dr. Hershman for persistent enthusiasm for this project and our collaboration.

I would also like to thank our other amazing collaborators, Dr. Kalinsky, and Dr. Lim, whose enthusiasm for this project has been a boon.

I would not have been able to do any of this research if it was not for the amazing team of women that I have worked with over the years. Our study saw three research coordinators: Mindy Brown, Susan Refice, and Jessica Campbell. Between scheduling, coordinating, and imaging patients, they helped keep this research going every week. I want to specifically thank Susan who was with us the longest stretch of the study and did such an amazing job.

Thank you to all the women who let me image them throughout the years and greatly contributed to this study. This thesis would not have been possible without their

time and commitment to breast cancer research. They have inspired me greatly throughout the years.

I would like to thank all my amazing lab mates: Molly, Ludguier, Michael, Chris, Jingfei, Jen, Martina, Lukas, Jonghwan, Allessandro, Willis, and Mirella. I need to thank Ludguier specifically, who was the one who convinced me to join Dr. Hielscher lab, when I had no idea what diffuse optical tomography was. I also need to thank Molly, who had been a patient lab mate and mentor for me when I knew so little about the field and was a great help when I was acclimating to the PhD environment. Additionally, she built the system that basically gave me my PhD thesis topic and for that I am ever so grateful. I need to thank Jonghwan who has helped me a lot during these five years, whether if it was finding the tools in the lab or listening to me ramble about possible ideas and give me feedback, he has been an incredible friend and colleague to have during this time. Also, thanks to Mirella, for these eight months of work that helped me keep on track and catch up with all the data this project had.

Thank you to all my friends who supported me constantly throughout my five year. Specifically, to Kelly and Tiffany who always believed I was capable of completing this overwhelming task. Thanks to Tala and Shih-ying for not only staying up late with me to finish homework assignments, but also imparting so much wisdom. I would like to thank all the people I have met at Intervarsity and their constant prayers for my work and research, especially Hon, Grace, Ariane, and Rubens. Also, I would like to thank CCM, Fr. Dan, and all my housemates at Ford Hall over the last two years: Gabriella, Chryshanie, Fran, Andy, Riccardo, Emily, Macrina, Christina, Paulette, Peter, Meli, Kevin,

Chris, Miguel, Kathy, Luisa, Shannon, Rebecca, Dylan, Elisa, Alyssa, Michelle, and Rosa. Thank you for your all your prayers and thank you for all the laughs.

I have to thank my parents who have constantly supported me throughout this entire journey and told me I could do this. I thank my sister Rebecca, who also supported me with both love and candy. I also need to thank the whole Gunther family for believing in me. I would like to specifically thank my grandfather, who was an engineer and has always been an inspiration to me. He told me when I was ten I should be an engineering and I stubbornly told him no. So I thank my grandfather, for believing I had what it took to become an engineer and do something great.

Finally, I need to thank the Good Lord for everything. In lumine Tuo videbimus lumen.

Dedication:

For my Lord

For my family

And for my friends

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

CHAPTER 2: STUDY DESIGN

CHAPTER 6: DYNAMIC RESULTS

CHAPTER 7: COMBINED ANALYSIS

CHAPTER 8: SUMMARY AND FUTURE WORK

LIST OF TABLES

CHAPTER 1: INTRODUCTION

CHAPTER 2: STUDY DESIGN

CHAPTER 3: STATIC PROCESSING

CHAPTER 4: STATIC RESULTS

CHAPTER 5: DYNAMIC PROCESSING

CHAPTER 6: DYNAMIC RESULTS

Acronyms

2D: two-dimensional
3D: three-dimensional
AC: alternating current
AC: doxorubicin and cyclophosphamide
ANOVA: analysis of variance
APD:
AUC: area under the curve
BC: breast cancer
BH: breath hold
BMI: body mass index
CC: correlation coefficient
CE: clinical exam
CW: continuous wave
DAQ: data acquisition
DC: direct current
DCIS: ductal carcinoma in situ
DCT: discrete cosine transform
DDOT: dynamic diffuse optical tomography
DITI: digital infrared thermal imaging
DOS: diffuse optical spectroscopy
DOT: diffuse optical tomography
EIS: electrical impedance scanning
EPI: electronic palpation imaging
ER: estrogen receptor
ES: evolution strategy
FD: frequency domain
FDG: fluorodeoxyglucose
FN: false negative
FP: false positive
Hb: deoxy-hemoglobin
[Hb]: deoxy-hemoglobin concentration
HbO_2: oxy-hemoglobin
[HbO_2]: oxy-hemoglobin concentration
HbT: total hemoglobin
[HbT]: total hemoglobin concentration
HER2: HER-2/neu
IE: initial enhancement
IFP: interstitial fluid pressure
ILC: invasive lobular carcinoma
ILC: invasive lobular carcinoma
LCIS: lobular carcinoma in situ
LDA: linear discriminant analysis
MG: x-ray mammography
MRI: magnetic resonance imaging
NACT: neoadjuvant chemotherapy
NMPV: normalized maximum peak value
NPV: negative predictive value
OI: optical imaging
pCR: pathological complete response
PDE: partial differential equation
PET: positron emission tomography
PIE: post-initial enhancement
PMT: photo multiplying tube
PPV: positive predictive value
PR: progesterone receptors
QDA: quadratic discriminant analysis
RCB: residual cancer burden
ROC: receiver operating curve
SD: source-detector
SNR: signal to noise ratio
SO_2%: oxygen saturation percentage
TD: time domain
TN: true negative
TP: true positive
US: ultrasound
USE: ultrasound elastography`
VOI: volume of interest
WF: water fraction
YI: Youden index

Publications

J. E. Gunther*, Emerson Lim*, Hyun Keol Kim, Molly Flexman, Jessica Campbell, Hanina Hibshoosh, Katherine Crew, Kevin Kalinsky, Dawn Hershman, and Andreas H. Hielscher "Dynamic diffuse optical tomography for monitoring neoadjuvant chemotherapy in breast cancer patients" (Submitted to Radiology)

Emerson Lim*, **J. E. Gunther***, Hyun Keol Kim, Molly Flexman, Hanina Hibshoosh, Kevin Kalinsky, Katherine Crew, Bret Taback, Jessica Campbell, Andreas H. Hielscher, and Dawn Hershman, "Diffuse optical tomography changes correlate with residual cancer burden after neoadjuvant chemotherapy in breast cancer patients." (Submitted to JAMA)

J. E. Gunther, Emerson Lim, Hyun Keol Kim, Molly Flexman, Kevin Kalinsky, Dawn Hershman, and Andreas H. Hielscher "Combination of static and dynamic optical tomographic imaging features to predict tumor response in breast cancer patients" (in preparation).

J. E. Gunther, Emerson Lim, Hyun Keol Kim, Molly Flexman, Kevin Kalinsky, Dawn Hershman, and Andreas H. Hielscher "Monitoring breast tumor progression during neoadjuvant chemotherapy using diffuse optical tomographic imaging" (in preparation).

M.L. Flexman, H.K. Kim, **J.E. Gunther**, E.A. Lim, M.C. Alvarez, E. Desperito, D.L. Hershman and A.H. Hielscher. "Optical biomarkers for breast cancer derived from dynamic diffuse optical tomography." *Breast Cancer Research. Journal of Biomedical Optics. 18(9)*, 096012, September 2013.

*Co-first authors

Presentations and Conference Proceedings

J.E. Gunther, E. Lim, H. Kim, M. Altoe, S. Refrice, M. Brown, K. Kalinsky, D. Hershman, A.H. Hielscher. "5-Month Monitoring of Tumor Response in Patients with Breast Cancer who undergo Neoadjuvent Therapy with Diffuse Optical Tomography" in *Biomedical Optics*, OSA Technical Digest (Optical Society of America, 2016) CM4A.2

J.E. Gunther, E. Lim, H. Kim, M. Flexman, L. Zweck, S. Arora, S. Refice, M. Brown, K. Kalinsky, D. Hershman, and A.H. Hielscher. Combined dynamic and static optical tomography for prediction of treatment outcome in breast cancer patients. Proc. SPIE 9538, Diffuse Optical Imaging V, 953811 (July 16, 2015); doi:10.1117/12.2183996.

J.E. Gunther, E. Lim, H. Kim, M. Brown, S. Refice, K. Kalinsky, D. Hershman, and A.H. Hielscher, Dynamic diffuse optical tomography for assessing changes of breast tumors during neoadjuvant chemotherapy. Proc. SPIE 9319, Optical Tomography and Spectroscopy of Tissue XI, 93190C (March 12, 2015); doi:10.1117/12.2079435.

Hyun K. Kim, **J.E. Gunther**, J. Hoi, A.H. Hielscher "A reduced-space basis function neural network method for diffuse optical tomography". Proc. SPIE 9319, Optical Tomography and Spectroscopy of Tissue XI, 931925 (March 5, 2015); doi:10.1117/12.2080550.

J. E. Gunther, E Lim, H. Kim, M. Flexman, M. Brown, S. Refice, K. Kalinsky, D. L. Hershman, and A.H. Hielscher. "Chemotherapy Monitoring of Breast Cancer Patients with a Dual-Breast Dynamic Optical Tomography System." in *Biomedical Optics*, OSA Technical Digest (Optical Society of America, 2014) BW4B.4.

E. Lim, **J.E. Gunther**, M. Flexman, H. Kim, H. Hibshoosh, K, Kalinsky, K. Crew, M. Maurer, B. Taback, S. Feldman, M. Brown, S. Refice, M.C. Alvarez, A.H. Hielscher, D.L. Hershman. **"**Two-Week Change in Optical Tomography Predicts Residual Cancer Burden Score in Women Treated with Neoadjuvant Chemotherapy." *San Antonio Breast Cancer Symposium.* Dec 2013.

J. E. Gunther, E Lim, H. Kim, M. Flexman, M. Brown, S. Refice, K. Kalinsky, D. L. Hershman, and A.H. Hielscher. "Predicting Tumor Response in Breast Cancer Patients Using Diffuse Optical Tomography." In Diffuse Optical Imaging IV, Proceedings of SPIE Vol. 8799 (SPIE, Bellingham, WA 2013), 87990P (June 14, 2013).

J. E. Gunther, E Lim, H. Kim, M. Flexman, M. Brown, S. Refice, K. Kalinsky, D. L. Hershman, and A.H. Hielscher. "A diffuse optical tomography imaging system for monitoring tumor response to neoadjuvant chemotherapy in breast cancer patients." *Photonics West.* Feb 2013.

J. E. Gunther, E Lim, H. Kim, M. Flexman, M. Brown, S. Refice, K. Kalinsky, D. L. Hershman, and A.H. Hielscher. "Diffuse Optical Tomography Imaging System for Neoadjuvant Chemotherapy Monitoring for Breast Cancer." *IEEE EMBS Healthcare Technology Conference* Nov 2012.

M. Flexman, **J. E. Gunther**, H. K. Kim, E. Lim, E. Desperito, R. L. Barbour, D. Hershman, and A. H. Hielscher, "Diffuse Optical Tomography Imaging of the Hemodynamic Response to a Breath Hold for Use in Breast Cancer Detection," in *Biomedical Optics*, OSA Technical Digest (Optical Society of America, 2012), paper JM3A.69.

J. E. Gunther, M. Flexman, E. Lim, H. K. Kim, M. Brown, D. Hershman, and A. H. Hielscher, "Diffuse Optical Tomography Imaging System for Monitoring Breast Tumor Response to Neoadjuvant Chemotherapy," in *Biomedical Optics*, OSA Technical Digest (Optical Society of America, 2012), paper BW3A.6.

M. Flexman, H.K. Kim, **J. Gunther**, E. Lim, E. Desperito, R. Barbour, D. Hershman, A.H. Hielscher, "Dynamic Optical Tomographic Breast Cancer Imaging." Poster presented at the Inverse Problems Workshop at Rensselaer Polytechnic Institute; 2011 October 27-28; Troy, New York.

CHAPTER 1: INTRODUCTION

1.1 Overall Goal and Specific Aims

The overall goal of this thesis is to determine whether optical tomographic imaging biomarkers can be employed to predict treatment outcome in women with breast cancer (BC) who undergo neoadjuvant chemotherapy (NACT). BC affects 1 in 8 women in their lifetime. In the United States, there are approximately 230,000 women with invasive BC and almost 40,000 die every year. NACT is widely applied as standard treatment for patients with newly diagnosed invasive BC. The goal of NACT is a pathologic complete response (pCR) at definitive surgery which typically is performed at ~6 months after treatment initiation. A pCR is associated with a substantial increase in 5-year survival and reduction in distant metastatic recurrences. However, a pCR is achieved in only 15-40% of women.[1-3] Furthermore, NACT is an expensive and toxic treatment that takes several months to completely administer, and major efforts are underway to find new drugs that increase pathologic response and ultimately improve survival. If the long-term treatment response could be predicted early into the treatment regimen, unnecessary side effects could be minimized, and alternative treatments could be initiated. [1, 3]

Diffuse optical tomography (DOT) has the potential for becoming an early outcome-prediction tool. Over the last decade, DOT has been shown to be very sensitive to vascular changes in BC and other diseases. As a result, several groups have started to study NACT monitoring using DOT. However, so far the patient-enrollment numbers have been relatively small and the few patients that participated in these studies have received a wide variety of different NACT regimens. Since the different drugs may affect the optical signals in different ways, the studies might have confounding results.

Furthermore, none of the previous studies have involved a full tomographic imaging system that images both breast at the same time, and very few studies have looked at hemodynamic changes in the tumor bearing breast.

To overcome the limitations of these previous studies, I had proposed to employ a novel dynamic DOT system to study the effects of targeted NACT. Unlike any other optical breast imager, this system provides full three-dimensional tomographic images of both breasts simultaneously. Furthermore, the system allows for the study of static and hemodynamic effects, which is achieved through relatively fast data acquisition frame rates. The goal was then to use the system to perform the largest clinical pilot study to date and only enroll patients that are on the same standard NACT regimen. To test and evaluate the potential of DOT imaging in the NACT setting, I have completed the following 4 aims.

<u>Specific Aim I</u>: Performed a 40-patient clinical NACT-monitoring study to demonstrate that 2-week changes in DOT-derived imaging biomarkers corresponds with pCR

Challenge: To determine if DOT can predict treatment response of breast tumors to NACT.

Approach: We have previously built and utilized a DOT system to image subjects with malignant and benign tumors, as well as, healthy breast tissue. We found that biomarkers such as oxy-hemoglobin [HbO_2], deoxy-hemoglobin [Hb], and total hemoglobin [HbT] concentrations, as well as their dynamic responses to a breath-hold, can be used to detect tumors. Employing the same system in this monitoring study, I hypothesized that these

biomarkers will change substantially at the beginning of the treatment and indicate if a patient will obtain a pCR.

Impact: Finding imaging biomarkers that can predict treatment outcome early on, will help physicians in designing more effective treatment strategies.

Specific Aim II: Developed a pre-processing and post-processing protocols for analyzing "static data" acquired from NACT imaging and to monitor how different chromophores change over the course of treatment.

Challenge: Artifacts at the tissue-air interface are very common in optical images and therefore can effect automated quantification. Additionally pre-processing and post-processing must be similar for all imaging time points.

Approach: Created custom algorithms via MatLab to pre-process data, to extract volume geometries that contain the tumor region, and to track the region of interest over time.

Impact: The static data shows absolute concentrations of hemoglobin and water fraction. Observations of these parameters show changes over time for subjects that have a good or partial response to treatment, and can be used to monitor treatment.

Specific Aim III: Develop a pre-processing and post-processing protocol for analyzing "dynamic data" acquired from NACT imaging.

Challenge: We can collect dynamic information over time, and are able to view the hemodynamic changes in the breast when a stimulus is applied (e.g. a 30-second breath hold). Therefore the challenge lies in how to use an intrinsic contrast to see tumor response.

Approach: Extracted features from dynamic breath hold images to determine different hemodynamic qualities of the vasculature of the tumor.

Impact: The dynamic optical imaging can show how blood flow changes in the tumor region during NACT.

<u>Specific Aim IV:</u> Combined static and dynamic data analysis schemes for discriminant analysis, which can be used to find a combination of imaging biomarkers that most accurately predict tumor response to NACT.

Challenge: Finding effective combination of physiological parameters that describe the response of the tumor and can distinguish response.

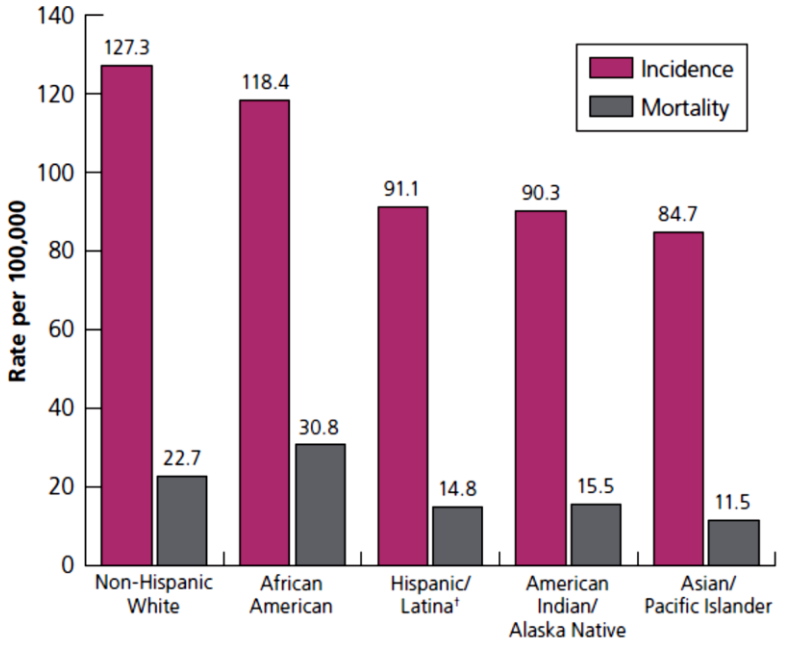
Approach: Used promising metrics from both the static and dynamic analyses to perform linear or quadratic discriminant analysis.

Impact: Determining a way to predict tumor response accurately using optical imaging.

1.2. Breast Cancer Epidemiology

The chance of a woman in the United States being diagnosed with breast cancer has increased to 1 in 8 (12.3%) from 1 in 11 women since the 1970's. The increase in prevalence is likely caused by several environmental factors that have change in the United States over the past 40 years. Some of these factors include the increase in obesity, increase life expectancy, changes in reproductive patterns (women having children at later ages), increase breast cancer screening, etc. Another factor for the increase in breast cancer cases may simply be the increase in breast cancer screening between 1979 and 1987. [4] The estimated number of incidences for 2014 is 232,670. The

number of new cancer cases has increased by 14% since 2000. Accordingly, the number deaths caused by breast cancer in 2014 has also increased and is estimated to be about 40,000 deaths or 6.8% of all cancer deaths.[5] However, as seen in Figure 1, the incidence and mortality rate can be very dependent on the race and ethnicity of the individual. There is usually a higher incidence of breast cancer in the non-Hispanic White population, yet the mortality rate is the highest among African Americans.[4]



*Rates are age adjusted to the 2000 US standard population.
†Persons of Hispanic origin may be any race.
Sources: Incidence: Copeland et al.[15] Mortality: Howlader et al.[14]
American Cancer Society, Surveillance and Health Services Research, 2013

Figure 1.1 Female Breast Cancer Incidence and Mortality Rates by Race* and Ethnicity, US, 2006-2010.[4]

The overall 5-year survival rate is about 89% among all populations in the United States.[4, 5] Yet, the 5-year survival rate can vary between 78.9% and 92.8% depending on race. As seen in Table 1.1, Asian women have the highest 5-year survival rate, while the African American population tends to have lowest survival rate. African American

women tend to be diagnosed with more aggressive cancers that have poor prognosis. Also, there are lower 5-year survival rates among populations that have low incomes, which prevent them from receiving the medical attention they require. Since in low income subgroups, there may be the lack of health care or funds to receive proper treatment.[4]

Table 1.1 Five Year Cause Specific Breast Cancer Survival by Race 2003-2009
*Survival rates are based on patient diagnosed between 2003 and 2009 and folled through 2010.[4]

Race	5-year Survival Rate
Non-Hispanic White	88.6%
African American	78.9%
Hispanic	87.0%
Asian	91.4%
Pacific Islander	86.8%

The stage of the cancer can affect the 5-year survival rate of the woman. Usually, when breast cancer is caught in its earlier stages, then the prognosis of the patient is much better. Therefore, as seen in Figure 2, localized cancer cases have the highest 5-year relative survival. While, distant or metastatic cancer usually has a poorer prognosis and a lower survival rate.[5]

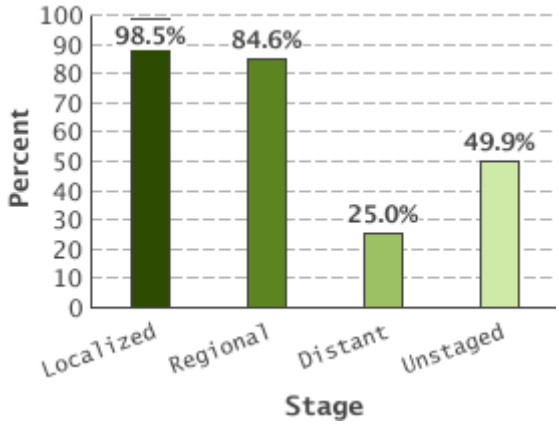


Figure 1.2 Five-year Relative Survival based on Stage of Disease[5]

Since the 1980's, the decrease in mortality rate has been contributed to increase screening using mammography (MG). However, in 2010, 51% of women over the age of 40 received a MG in the last year and 61% of women received a MG in the past two years. Additionally, this percentage changes depending on income and age. For non-poor women between 74%-83% received a MG within the last two years. While only 46%-57% of women who are poor or near-poor received a MG in the past two years.[4] Therefore, there is a clinical need to decrease the cost BC screening in order to increase the number of women who routine screening.

There are other factors that can prevent BC screening. I have spoken to women over the past 5 years about the process of breast cancer screening and treatment and many have communicated their dislike of various imaging modalities. Some women refuse to get MG imaging done because the compression is too painful. Other women have said that they are afraid that the radiation will cause breast cancer in the future. Some women have refused MRI because they are uncomfortable or claustrophobic. Despite the risk of a having BC, many women still refuse to undergo the recommended procedures for tumor detection or monitoring. Hence, BC imaging has other facets that need improvement to encourage women to get routine screening.

1.3 Breast Cancer Pathology

Breast tissue is mostly consistent of adipose, fibrous, and glandular tissue. Each breast has about 15-20 lobes, which each has up to 40 lobules.[6, 7] The lobules are further divided into 10-100 alveoli, which secrete milk. The alveoli are connected to a lactiferous tubule, which is connected to a lactiferous duct. These ducts are basically conduits that

bring milk from the milk producing cells to the nipple. Last, the glandular complexes are surrounded by adipose tissue.[6]

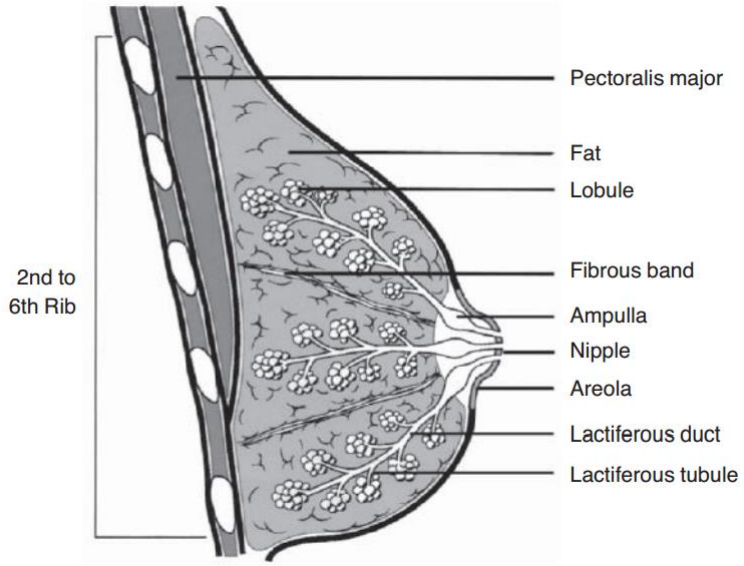


Figure 1.3 Basic anatomy of the breast.[6]

Breast cancer typically originates in either the lobules or the ducts of the breast. The four most common types of breast cancer are invasive ductal carcinoma (IDC), invasive lobular carcinoma (ILC), ductal carcinoma *in situ* (DCIS), and lobular carcinoma *in situ* (LCIS). DCIS originates in the duct of the breast, while LCIS originates in the lobules. *In situ* type cancers arises when abnormal cells grow and proliferate within the ducts or lobules without invading the basement membrane.[8] IDC occurs when tumor cells have invaded the basement membrane. Therefore DCIS is often a precursor for IDC.[8, 9] Similarly, ILC occurs when infiltrates the stroma of the breast and happens to be much more difficult to detect.[8]

Cancer, in general, starts in an avascular stage where the tumor will grow to between 1-2 mm and then reach an equilibrium of growth and apoptosis. Tumors cannot

grow more than a couple millimeters without some conduit to receive nutrients. Hence, tumors need to recruit blood vessels in order to progress. Hypoxia in the tumor region can also induce these new blood vessels.[10]

Tumor angiogenesis is much different from normal physiological angiogenesis. Tumor blood vessels are irregularly shaped, dilated, have dead ends, and have increased permeability.[10] Also, blood vessels in the tumor are poorly organized, are compressed by surrounding tumor tissue, and provide limited delivery of oxygen and nutrients, and have irregular blow flow.[11] Additionally, the blood vessels are forced apart due to the growing tumor, which can prevent proper drug penetration.[11]

1.4 Breast Cancer Imaging

There are several different methods for imaging breast tissue and identifying breast cancer. Traditional methods include X-ray mammography (MG), ultrasound (US), magnetic resonance tomography (MRI), and positron emission tomography (PET). However, there have been experimental imaging modalities that have been developed, such as optical imaging (OI). To better understand how OI can be utilized in the clinic, we must examine the imaging modalities that currently exist and how these methods are being utilized in the breast cancer setting.

1.4.1 X-ray Mammography

Over the last 25 years, the gold standard for breast cancer detection has been X-ray mammography (MG). In the late nineties, there have been some reports of a ~25% reduction in breast-cancer mortality; More recent studies done in 2012 showed a reduction between 7%-23% in Europe and Canada over the last couple of decades.[12, 13]

In general this drop has been attributed to an increase in screening over the previous decades. However, it should be noted that during the same time period there has also been an increase in adjuvant chemotherapies which, some argue, may have contributed to two thirds of the mortality rate decrease.[14] Yet another major study done in Canada suggested that MG has no effect on the mortality rate, and overall there seems to be no clear consensus about the advantages and disadvantages of X-Ray imaging in breast cancer detection .[15]

While there are conflicting results about the effect of MG on the mortality of breast cancer patients, there is no doubt about the diagnostic capabilities of X-ray imaging. According to the United States screening program MG has a sensitivity of 75%. The sensitivity decreases to 30%-40% with a false negative (FN) rate of 4%-34% for women with denser breasts. [12, 13] Therefore, MG is currently suggested for women over the age of 40 years, since the sensitivity of MG increases as breast density decreases with age and a postmenopausal status.[12, 16]

The general guidelines suggest that women between the ages of 40-49 receive biennial MG screenings, while women between the ages of 50-74 should receive annual screenings. [12, 13] In the state of New York screening coverage laws allow baseline imaging at 35-39 years of age and every year over the age of 40 or with a physician recommendation. Insurance coverage for MG screening varies from state to state within the United States.[17]

1.4.2 Ultrasound

In general, US is not recommended as a screening tool and is usually not covered by insurance. Yet, US is employed as an adjunct imaging method to MG. Also US is

considered if the patient is at high risk and cannot tolerate MRI, or if they have an intermediate risk with dense parenchyma or dense breasts. High resolution US can also be used to detect occult cancers or be useful with patients with dense breasts. However, there is limited detection of DCIS and microcalcifications. Additionally, a negative US never supersedes the results of a MG, and negative US result does not indicate the lack of a tumor.

Studies have shown that US detects 2.7-4.6 cancers per 1000 asymptomatic women who have dense breast or negative MG findings. Yet, the use of US imaging leads to a high false positive rate and a low positive predictive value (PPV) of 8.8%-11%.[12]

A 2002 study looking at 27,825 patient examinations (11,130 asymptomatic women) found 246 cancers in 221 women. The addition of US screening allowed for the detection of smaller and lower stages of cancer than physical examination (PE) and increase detection in occult cancer. The sensitivities for the study were 77.6%, 27%, and 75.3% for MG, PE, and US, respectively. Sensitivity of US was more or less unaffected by breast density or age. From the women with non-palpable invasive cancer, 42% of cancers were detected with US only. From the subgroup of US only detected cancer 70% were <1 cm and 89% were node negative, which would give the best prognosis and the largest number of options for treatment. However, the authors mention that the benefits of the screening must be weighed against the false positive (FP) rate of 2.4%. Also rate of positive biopsy due to US detection is much lower than in MG.[16]

Another study done in Germany in 2012 suggests that there is little or marginal benefit to US screening. They were able to detect additional cancers that MG did not, but the PPV for the study was 0.21%.[18]

Therefore, US seems has some benefits as a screening tool, but at the risk of additional FPs and unnecessary biopsies. Additionally, US is highly operator dependent.[12, 16] However, US has been useful as a follow-up imaging modality for patients with dense breasts or at high risk.

1.4.3 Magnetic Resonance Imaging

MRI has been used for staging and treatment planning and has shown to be helpful at screening high risk patients.[12, 13, 19] As seen in Table 1.2, there are different recommendations for screening depending on the cancer risk of the patient. Women who are at high risk are encouraged to get annual MG and MRI imaging done. Annual MRI screening can start as early as age 30. High-risk patients usually are carriers of deleterious BRCA genes, have first-degree relatives with BRCA mutation, or received chest irradiation between 10-30 years of age. [12, 13] High-risk women tend to have an earlier onset of cancer and therefore, may benefit from earlier screening using MRI, since MG would have lower sensitivity in patients under the age of 40. MRI is usually not recommended for average risk patients even if they have dense breasts.[12]

MRI shows higher sensitive of over 90%. Yet, there is still a low specificity of 40%-80%. Sensitivity tends to be greater in invasive carcinomas compared to DCIS. When combined with MG there is a 58% increase in sensitivity and a 44% increase when combined with MG and US compared to MG alone.[13]

Table 1. 2 Recommended screening for patients depending on risk of cancer.[13]

Risk	Population	Recommendation
Average risk	Women with <15% lifetime risk for breast cancer	Annual MG starting at age 40
High risk	BRCA1 or BRCA2 carriers, untested first degree relatives of BRCA mutation carriers	Annual MG and MRI starting by age 30 but not before age 25
	Women with >20% lifetime risk for breast cancer on the basis of family history	Annual MG and MRI starting by age 30 but not before the age 25, or 10 years before the age of the youngest affected relative, whichever is later.
	History of chest irradiation received between the ages of 10 and 30	Annual MG and MRI starting 8 years after treatment; MG is not recommended before age 25.
Intermediated risk	Personal history of breast cancer, ovarian cancer, or biopsy diagnosis of lobular neoplasia or atypical ductal hyperplasia	Annual MG from time of diagnosis; either annual MRI or US can also be considered
	Women with dense breasts as the only risk factor	The addition of US to screening MG may be useful for incremental cancer detection.

Overall, MRI can detect cancers that are not visible via other modalities due to its higher sensitivity, especially DCIS and ILC. Although MRI can be used for screening high risk patients, it has also been utilized for staging and treatment planning for new diagnosed patients. However, preoperative or perioperative MRI may not be beneficial for patient care since it can lead to a higher false positive rate, delay in treatment, higher costs. Also, there is no significant effect on the long-term survival of the patients due to MRI results compared to standard imaging methods. Additionally, there is evidence in the literature to suggest that perioperative MRI has led to an increase in mastectomies.[19]

1.4.4 Positron Emission Tomography

PET is a non-invasive exploratory technique for breast imaging, which uses radioactive tracers such as fluorodeoxyglucose (FDG) to obtain information about the uptake and metabolism of glucose. Malignant tissue tends to have both a higher uptake of glucose and a fast metabolism, and allows for detection via PET. For detecting primary

tumors FDG-PET, the review article reports a range of 48-95.7%. However, sensitivity reduces when the tumor is sub-centimeter due to a decrease uptake in FDG. FDG-PET seems to be a more reliable modality compared to conventional techniques (MG, US, etc.), but does not out preform magnetic resonance mammography and MRI when diagnosing primary tumors. Yet, FDG-PET can only be recommended for the detection of metastases and recurrent disease. Compared to conventional imaging, FDG-PET shows increased efficacy and similar sensitivity, but with lower false positives.[20]

1.4.5 Experimental Imaging Modalities

There are additional imaging modalities that are still within the experimental stages and have not been widely used in a clinical setting yet, according to Vreugdenburg, et. al. A few of these imaging modalities include digital infrared thermal imaging (DITI), electrical impedance scanning (EIS), and elastography, which include electronic palpation imaging (EPI) and ultrasound elastography (USE).

DITI produces a heat map or thermogram of the breast and is captured using infrared cameras. Usually, DITI has a very low sensitivity, but with development of high resolution infrared cameras the topic has been revisited. Tumors can be detected by DITI due to localized increase in skin temperature, since tumor display increased vascularity, vasodilation, and recruitment of inflammatory cells.

Cancer cells have increased electrical conductivity and lower impedance compared to healthy cells. Therefore, EIS has also been utilized to locate breast tumors.

EPI creates pressure maps of the breast where the tumor is a more rigid structure compared to healthy tissue. USE compares the strain between healthy and cancerous tissue.

However, for this review article, the authors mention that a few studies have been done on asymptomatic patients and therefore they do not recommend including these particular imaging modalities as screening tools for the time being.[20]

Table 1.3 Summary of the sensitivity and specificity of DITI, EIS, and USE from literary search. [20]

Modality	Sensitivity (Median)	Specificity (Median)
DITI	25.0-96.7% (82%)	11.8-84.9% (55%)
EIS	62.0-97.5% (83%)	42.0-80.9% (68%)
USE	35.4-100% (84%)	21.1-98.9% (88%)

1.4.6 Problems with Current Imaging Modalities for Breast Cancer

Despite the multiple imaging modalities available and years of screening studies done in the field of breast cancer detection, there is still room for improvement. The overall goals are to reduce the rate of late-stage disease, decrease cancer mortality, and detect aggressive tumors early. As mentioned previously, MG tends to lead to FPs, which leads to the problem of overdiagnosis. Overdiagnosis is the detection of cancer that if left untreated would not be clinically significant or detrimental. As a result, overdiagnosis can then lead to overtreatment. [14, 21]

In a Canadian study involving 89,835 women being screened for breast cancer, there was one over-diagnosed breast cancer for every 424 patient who received MG imaging. Additionally they found that annual MG for women 40-59 did not reduce the mortality rate. As seen in Figure 1.4, there is very little difference between the survival probability between the group of women that received MG screening and the control group that did not receive imaging. [15] The results found in this study seem contrary to what has been reported in the literature previously. Nevertheless, the report has sparked

media attention in the United States, and has women asking if MG is necessary or worthwhile.

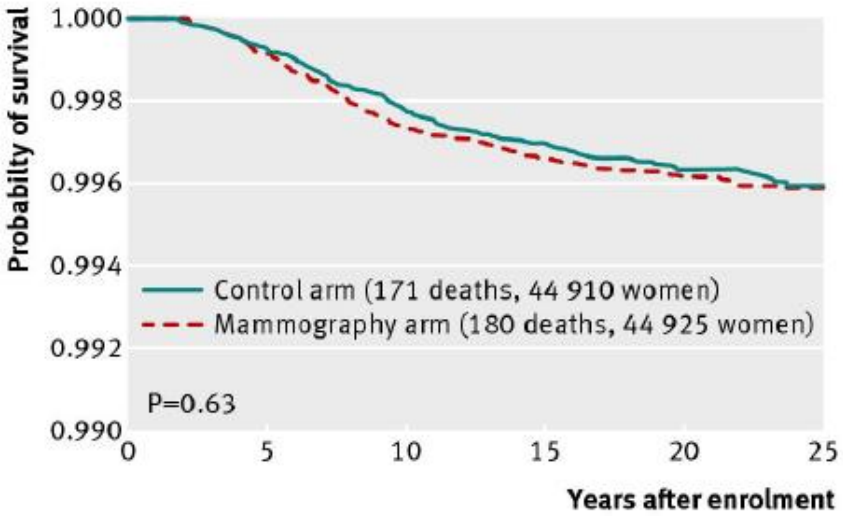


Figure 1.4 Breast cancer specific mortality from cancers diagnosed in screening period, by assignment to mammography or control arms.[15]

Survival rates for breast cancer have improved over the past few decades. However, the reason behind this trend can be debated. The decrease in mortality rate may be due mostly to the development of adjuvant therapies and better treatment plans for patients and not so much the early detection of cancer. Therefore, there is a need for early detection of aggressive tumors before it progresses or form metastases. However, there is the issue of the preconceived notion among doctors and patients of not doing enough for a patient even if the cancer is indolent.[21] Additionally, there is a need for imaging technology that can differentiate between benign and malignant tumors.

1.5 Optical Imaging

Since there are areas of breast cancer imaging that need improvement, OI imaging methods, such as diffuse optical tomography (DOT), have emerged as a novel imaging

modality for this setting. DOT uses optical transmission measurements with non-ionizing, low-intensity, near-infrared light to generate three-dimensional maps of various tissue parameters. Optical property contrast from endogenous chromophores [oxyhemoglobin (HbO_2), deoxyhemoglobin (Hb), water, and lipid] can distinguish malignant from normal tissue.[22-25] For example, it has been reported that total hemoglobin (HbT) levels, which relate to tumor blood vessel density, are double those in benign breast lesions[26]; and vascular changes precede measurable structural changes in mouse models.[27]

DOT works on the principles that govern light interaction in biological tissue. Different types of optical imaging systems can be employed to capture this light interaction. Afterwards, light propagation models and optimization algorithms can be utilized to reconstruct the optical properties and chromophore concentration in the tissue.

1.5.1 Light and Tissue Interactions

Biological tissue is optically strongly scattering and tends to be weakly absorbing especially within the 400-1350 nm spectral region. Absorption and scattering are light phenomena that occur when light is shined through a medium. Absorption refers to a decrease in intensity of light as it travels a certain path distance (x) as shown by equation 1.1.

$$I(x) = I_0 e^{-\mu_a x} \tag{1.1}$$

I(x) is the intensity of light after it has travelled distance x. I_0 is the initial intensity of light and μ_a is the absorption coefficient. Major causes of absorption in biological tissue come from hemoglobin, melanin, and water.[28] Figure 1.5 shows oxy- and deoxy-hemoglobin absorption spectrum in the near-infrared range.

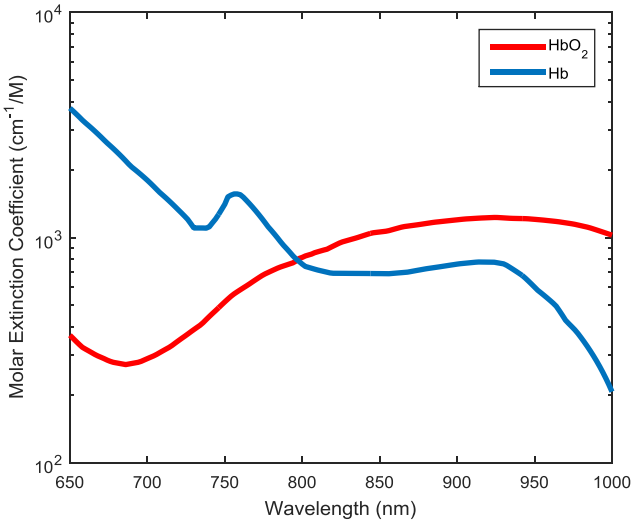


Figure 1.5 The molar extinction coefficients for oxyhemoglobin (HbO2) and deoxy-hemoglobin (Hb) for light in the near infrared range.[29]

Scattering is when light changes direction due to some obstacle or media of different optical properties. The scattering coefficient μ_s is the probability that a photon will scatter per unit length. Beer's law describes the probability of transmittance (T) through distance x (Equation 1.2).

$$T(x) = e^{-\mu_s x} \tag{1.2}$$

Scattering in biological tissue usually occurs when light interacts with structures such as nuclei.[28]

1.5.2 Optical Imaging Systems

After understanding light interaction in biological tissue, optical systems can be developed to capture this phenomenon. There are three types of optical systems: time domain (TD), frequency domain (FD), and continuous wave (CW). Table 1.4 shows a summary of the different optical systems.

Table 1.4 Summary of types optical systems.[28]

Type of Imaging System	*Source Light*	*Detection*	*Design*
Time Domain	Impulse	Photon counter	Information-rich but slow DAQ and most expensive
Frequency Domain	Amplitude-modulated	Amplitude and phase of emitted light	Faster and less expensive than TD.
Continuous Wave (Steady State)	DC	Amplitude	Fast DAQ Dynamic imaging Least Expensive

TD systems work by sending short light pulses through a medium and detecting the light as a function of time using a photon counter.[28, 30, 31] The resulting signal has a lower peak amplitude and has a broader curve than the input signal. The peak of the detected signal yields information about the path length of the light. A highly scattering medium would have a later peak due to the longer path length. The decreasing slope of the signal gives information about the absorption of the medium. A sharper slope would mean higher absorption.[30]

FD systems use lasers that modulate light in the 100-1000 MHz range. The system detects the amplitude and the phase shift of the signal, which can be used to determine absorption and scatter. The same information as TD systems is acquired, but FD systems cost much less.[28, 30, 31]

Last, CW systems use unmodulated or low modulated (<10 kHz) light that detects only the amplitude of the signal. Some of the advantages are that these systems are low cost and can be used for dynamic imaging.[28, 30, 31] However, it is difficult to distinguish absorption from scattering effects by using this approach.[31]

Table 1.5 Summary of clinical optical breast imagers.

Optical Group	Modality	Method	Acquisition Type	Sources	Detectors	wavelengths [nm]	Modulation Frequency	Total Imaging Time
Boas[32] MGH	Tomographic optical breast imaging system. Combined Optical and Mammography; X-ray used for mesh generation	Compressed breast between source (20 cm x 18 cm) and detector probe	1st CW	3 frequency-encoded lasers - Galvo scanner - 110 points (200ms per point)	32 APD	685 nm, 810 nm, 830 nm		45 s for more than one source scan
			2nd CW	26 frequency-encoded lasers		685 nm, 830 nm		
			FD	40	8 APD	685 nm, 830 nm	**70 MHz**	
Zhu[33] UConn	Combined US and Optical system	Handheld probe	FD	1st system: 12	8 parallel PMT	780 nm, 830 nm	**- 50 MHz - 140 MHz - 350 MHz**	5 to 8 minutes for 10 to 20 data sets at a lesion site
				2nd system: 9	10 parallel PMT	prototype: 690 nm, 780 nm 830 nm		
Pogue[34] (Dartmouth)	Combined Optics and MRI system	open breast coil inside the MRI	FD	16 (6 laser diodes)	15 parallel PMT	661nm 752 nm, 785 nm, 805 nm, 829 nm, 849 nm	**100 MHz**	measurement with 6 wavelengths takes approximately 4 min
Tromberg[35] (UC Irvine)	Combined FD and CW system	Handheld probe; source to detector: 21.5mm (FD) and 24mm (CW)	FD	7 amplitude-modulated diode lasers	8–12 min (for one breast)	672 nm, 800 nm, 806 nm, 852 nm, 896 nm, 913 nm, and 978 nm	251 modulation frequencies from **100 to 700 MHz**	40 s per sample (30 s for FD and 10 s for CW); Take 20 minutes to 1 hours to image
			CW	150-W halogen lamp	fiber-coupled spectrograph with a linear CCD detector from 525 to 1155 nm	650–1000 nm		
Yodh[36] (UPenn)	FD and CW system	parallel-plate (“slight compression”) built-in Intralid/Ink fluid tank	FD	45 sources (**4 laser** diodes); 9 x 5 grid pattern	3 x 3 grid (FD and CW-mode)	690 nm, 750 nm, 786 nm, and 830 nm	**70 MHz**	8-12 min (for one breast)
			CW	45 sources (**2 laser** diodes); 9 x 5 grid pattern	lens-coupled 16-bit CCD camera (41 x 24 grid)	650 nm, 905 nm		

Table 1.5 shows a summary of clinical optical imagers from several optical imaging groups that are investigating breast cancer in the clinical setting. As for the groups led by Boas, Zhu, and Pogue, they combine traditional imaging modalities with optical imaging (MG, US, and MRI).[32-34] Tromberg's system is a DOS system that has one source and one detector. The DOS system is used to create line scans of the tumor region, which can be used to create 2D surface maps of optical parameters.[35] Last, Yodh's group does DOT imaging with a parallel plate design and is capable of creating 3D tomographic images.[36]

1.5.3 Iterative Image Reconstruction Algorithm

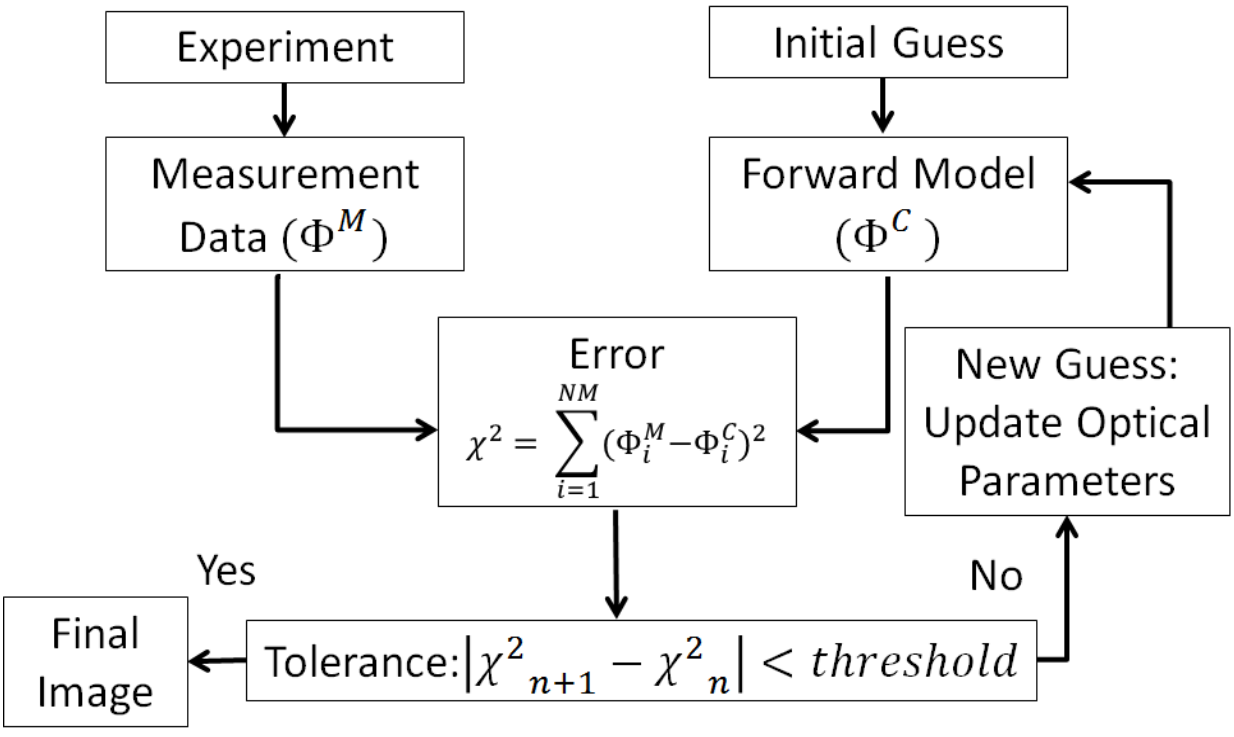


Figure 1.6 Flow diagram of the optimization algorithm for the inverse problem.[30]

After the data has been collected the images are reconstructed. The first component of the image reconstruction algorithm is the initial guess, which is an estimate of the spatial parameters of μ_a and μ_s (see Figure 1.6). The forward model is run to predict how light will travel through the medium with the parameters from the initial guess. From there the prediction model is compared to the experimental data and the error between these two measurements is found. If the error is not below a certain threshold the optical parameters are changed and updated and the forward model run again. The updating phase of the algorithm takes place until the predicted model and experimental measurements are less than a designated threshold in which the final image is presented.[30]

1.5.4 Light Propagation Models

In order to run the iterative scheme, there needs to be a forward model in place. The forward model is the light propagation model that predicts how light will travel through a medium of certain optical properties. The most basic form of light propagation through a medium is the radiative transform equation (RTE) and is seen in equation 1.4.

$$\frac{\delta\Phi(\vec{r},\hat{s},t)}{\delta t}\frac{1}{c} = -\hat{s}\cdot\nabla\Phi(\vec{r},\hat{s},t) - \mu_t\Phi(\vec{r},\hat{s},t) + \mu_s\int_{4\pi}\Phi(\vec{r},\hat{s},t)p(\hat{s}'\cdot\hat{s})d\Omega' + S(\vec{r},\hat{s},t) \qquad (1.3)$$

The spectral radiance, Φ, is the energy rate per area per solid angle (W m^{-2} sr^{-1}). The spectral radiance is terms of position (r), the unit direction vector (s), and time (t). The source term is denoted by S and μ_t is the summation of μ_a and μ_s'. The function S is the spectral radiance of the source. The function, p, is the probability that light will probability of photons scattering from $\hat{s}_{in}$ to $\hat{s}_{out}$ and is defined by the Henyey-Greenstein phase function:

$$p(\hat{s}_{in}\cdot\hat{s}_{out}) = \frac{1}{4\pi}\frac{1-g^2}{(1+g^2-2gcos\theta)^{2/3}} \qquad (1.4)$$

where g is anisotropy. Anisotropy can range from -1 to 1 with a value of zero would indicate isotropy. For most biological tissue g=0.9, which indicates that the medium is highly forward scattering.

However, for light modelling in biological tissue such as breast tissue the diffusion approximation is used since the medium is highly scattering. The diffusion approximation is derived from the RTE. To obtain the diffusion approximation there are a two of assumptions made:

1. The medium is nearly isotropic
2. $\mu_a << \mu_s$.

After extensive derivation the diffusion approximation is:

$$\frac{\delta I(\vec{r},t)}{\delta t}\frac{1}{c} + \mu_a I(\vec{r},t) - \nabla \cdot [D\nabla I(\vec{r},t)] = S(\vec{r},t) \quad (1.5)$$

If we are concerned with steady state where $\frac{\delta I(\vec{r},t)}{\delta t} = 0$ then the diffusion approximation is:

$$\mu_a I(\vec{r},t) - \nabla \cdot [D\nabla I(\vec{r},t)] = q(\vec{r},t) \quad (1.6)$$

where I is the intensity (W m^{-2}) and can also be defined as:

$$I(\vec{r},t) = \int_{4\pi} \Phi(\vec{r},\hat{s},t)\, d\Omega \quad (1.7)$$

Since I has no dependency on angle, the light is isotropic. The source term expressed as an intensity is q. D is the diffusion coefficient and is calculated using:

$$D = \frac{1}{3(\mu_a+\mu_s')} \quad (1.8)$$

In a discretized space, each unit (voxel) has designated optical properties (μ_a, μ_s') for each wavelength being observed. When the reconstruction algorithm updates, it is updating the optical properties of each unit to closer mimic the optical properties from the imaged volume. After the updating the optical properties, this information is used again

to determine how the light travels through the medium and what the intensity will be on the surface of the medium where the measurements take place. Therefore, optical properties such as μ_a, μ_s, and D are vectors that denote value for each voxel. Essentially, the reconstruction will develop a three-dimensional map of the optical properties.

1.5.5 Chromophore Concentration

After the optical properties are determined using the reconstruction algorithm, the concentration of a chromophore can be found using Beer-Lamberts Law, which is defined as:

$$\mu_a(\lambda) = \sum_{i=1}^{N} \varepsilon_i^{\lambda} C_i \quad (1.9)$$

where λ is the wavelength being employed, ε is the extinction coefficient, and C denotes the concentration.[28]

1.6 DOT Breast Imaging

DOT uses optical transmission measurements with non-ionizing, low-intensity near-infrared light to generate three-dimensional maps of various tissue parameters. Optical property contrast from endogenous chromophores can distinguish malignant from normal tissue[25, 37-39]. For example, it has been reported that total hemoglobin (HbT) levels, which relate to tumor blood vessel density, are double those in benign breast lesions[26]; and vascular changes precede measurable structural changes in mouse models.[27]

In a systematic review, Leff et. al. looked at several studies that involved 2,000 women in total and found that about 85% of tumors were detectable by using optical techniques. They determined that hemoglobin concentration and oxygen saturation were

the two important factors for determining breast disease. Combining the studies they found that the average [HbT] for malignant disease was 65±35 μmol/l and SO_2% was 66±24%. For fibroadenomas the average was 54±13 μmol/l and 69 ± 3% for [HbT] and SO_2%, respectively. In healthy tissue, [HbT] was 21 ± 6 μmol/l and SO_2% was 68 ± 5.[31] The [HbT] for malignant disease were much higher than healthy tissue, while SO_2% had similar levels. However, there are still some difficulties at distinguishing between malignant and benign tumors. The best results for tumor differentiation seem to be with FD systems with US or MRI components.[31]

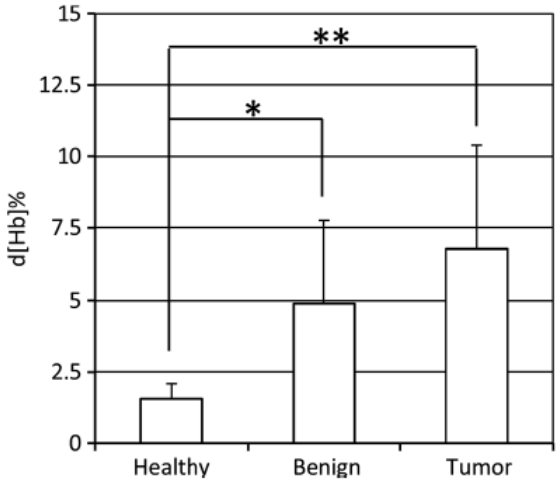


Figure 1.7 Statistical analysis of the differences in Δ[Hb]% between healthy (n = 6), benign (n = 8), and tumor-bearing (n =14) breasts 15 seconds after the end of a breath-hold. (a) Mean value and standard deviations of the average Δ[Hb]% in a 1-cm-sphere around the point of peak hemoglobin change. [40]

In a previous study, our group at Columbia University used dynamic diffuse optical tomography (DOT) imaging to observe the hemodynamics of different tumors to determine if a subject had a benign or malignant tumor, or was healthy. The contrast mechanism that was used for the pilot and current study was a simple breath hold. When a subject holds their breath there is an increase in intrathoracic pressure, which decreases the venous return of blood to the heart. Hence, there is an increase in blood

volume in the breast and when the subject breathes normally the blood volume returns to normal. The quantified hemodynamic time traces showed variation between the healthy and tumor bearing breasts of a subject compared to that of a healthy subject who showed similar traces for both breast. Additionally, there were slower recovery times for the tumor-bearing breast and higher Δ[Hb]% at the 15 second post breath hold image as seen in Figure 1.7 [40]

1.7 Monitoring Neoadjuvant Chemotherapy

1.7.1Monitoring NACT using Traditional Imaging

NACT is widely applied as standard treatment for patients with newly diagnosed operable invasive BC.[41] The goal of NACT is shrink the breast tumor as much as possible before surgical intervention. In this way an inoperable cancer may become operable, or an operable cancer can be converted from mastectomy to breast conserving therapy. Ideally a pathological complete response (pCR), meaning the complete disappearance of the tumor before surgery, is achieved. Studies have shown that patients with a pCR have a 90-95% five-year survival rate, while five-year survival rates for patients with no pCR are typically below 80%.[42] Depending upon the exact definition, pCR is achieved in only 15-40% of women. Recent studies have shown that changing therapy in patients that do not show an early response may benefit from a change in treatment regimen.[43-47] Therefore, it would be highly desirable to reliably determine who will and will not respond to NACT early in the treatment.[3, 48]

Most NACT monitoring studies have concentrated on baseline and pre-surgical imaging to observe any residual cancer and estimate the size of the remaining tumor.

This is more or less for surgical planning purposes to assist physicians on whether breast conserving surgery is the best option or preform a full mastectomy. The goal is to predict which patients received a pathological complete response and prevent surgery all together. However, none of the traditional methods, such as clinical exam (CE), MG, and US seem to be accurate enough to estimate the pathological response or size of the residual tumor.[2, 49-54] Some imaging studies have shown that MRI is more accurate than other imaging methods[50, 51], yet other studies have seen little difference or have not been able to reliable predict response.[2, 53] A meta-analysis of forty-four studies with 2050 patients showed the MRI had a higher accuracy than CE or MG, but no difference between MRI and US. The median sensitivity from the meta-analysis for MRI was 92% and the median specificity was 60%.[54]

Croshaw et. al. combined retrospective data with 6 other studies to determine the efficacy of CE, MG, US, and MRI, as seen in Table 1.6. US has the best sensitivity of all the imaging modalities, but consequently the lowest specificity, which was the result of the highest number of false positives. CE had the lowest sensitivity at 50% but the highest specificity of 82%. Last, MRI showed the highest accuracy when determining pCR.[53]

Table 1.6 Results of Data Synthesis to Predict Tumor Response.[53]

	CE	Digital MG	US	MRI
Total Patients	259	255	301	230
Sensitivity	50%	81%	90%	86%
Specificity	82%	48%	33%	79%
Accuracy	57%	74%	79%	84%

One study did imaging at the mid-treatment time point to see if tumor response to NACT could be predicted before surgery. When predicting the response of the tumor at the mid-time point the accuracy of CE, MG, US, and MRI was 58%, 56%, 63%, and 67%, respectively.[51] Not only do none of these methods seem better than the others including

CE, but also none of these methods have a high enough accuracy to be clinically useful in predicting response at the mid-time point.

1.7.2 Optical Tomographic Imaging for Neoadjuvant Chemotherapy

Although there have been several tumor diagnosis OI and DOT studies[25, 26, 36, 55, 56] including our own[23] that show promising results, the technology has found a niche in NACT monitoring.[3] Since traditional method lack the accuracy or the feasibility for treatment monitoring or predicting response, DOT has emerged as a promising novel imaging modality for this important application

There are several advantages to using optical imaging (OI) to monitor breast tumor response over conventional imaging techniques. For example, OI using non-ionizing light sources that allow for multiple imaging session within a short period of time. Therefore, measurements can be taken a few weeks to a few days apart from each other with no adverse effect. OI is also sensitive to parameters that change over the course of treatment, such as tissue vascularity or hypoxia.[3, 48] Several studies have shown that there is a change in optical properties over time during the course of neoadjuvant chemotherapy. [48, 57-60] The purpose of these monitoring studies is to determine early within the treatment whether the patient will respond to the therapy or not. Therefore, more personalized treatments could be developed to optimize the patient outcome of those who do not respond. [3, 48]

Some studies have demonstrated that there is a decrease in total hemoglobin concentrations ([HbT]) time for subjects that had received a pathological complete response (pCR) or near-pCR.[57, 58] Another study using spectroscopic methods examined 34 women and found significant difference mid-treatment time point between pCR and

non-pCR subjects where the pCR subjects had a larger decrease in normalized deoxyhemoglobin concentrations ([Hb]).[59] Additionally, the same group was able to predict tumor response to NACT using the baseline measurements and determined that oxygen saturation was higher in patients that were going to receive a pCR.[60]

Since optical imaging has been shown to predict tumor response we performed a clinical study monitoring tumor response using our dynamic optical imaging system.

1.8 Overview of this Thesis

This work aimed to employ an existing dynamic diffuse optical imaging system to image breast cancer patients undergoing NACT. For the study, two different types of data were collected: static data (single 3D image) and dynamic data (3D movies). The goal was to find imaging biomarkers that could be used to accurately predict treatment response of the patient as early as two weeks after treatment initiation. Currently physicians need to wait until the 5 months of NACT has been completed and the patient had undergone surgery, to find out if a patient had a good response to treatment. If not, it is often too late to start a new therapy. With non-invasive, non-ionizing DOT imaging we may be able to determine within weeks of treatment if the tumor will disappear in 5 months; giving the physician and patient new options in therapy before expensive drugs with considerable side effects have been used for month.

Chapter 2: Study Design

2.1 Introduction

Neoadjuvant chemotherapy (NACT) is widely applied as standard treatment for patients with newly diagnosed operable invasive BC.[41] The goal of NACT is to shrink the size of the breast tumor as much as possible before surgical intervention. In this way an inoperable tumor may become operable, or breast conserving surgery can be performed instead of a mastectomy. The best outcome would be a pathological complete response (pCR), in which there is no residual cancer after treatment. Studies have shown that patients with a pCR have a 90-95% five-year survival rate, while five-year survival rates for patients with no pCR are typically below 80%.[42] Depending upon the exact definition, pCR is achieved in only 15-40% of women. Recent studies have shown that changing therapy in patients that do not show an early response may benefit from a change in treatment regimen.[43-47] Therefore, it would be highly desirable to reliably determine early in the treatment, who will and will not respond to NACT.[3, 48]

Our NACT monitoring study consists of imaging 40 women at six different time points during their treatment. We have hypothesized that DOT imaging can determine early within treatment whether the patient will or will not have a pathological response to the chemotherapy. For this chapter, we provide the methods and overview of the clinical pilot study.

2.2 Clinical Study Overview

2.2.1 Patient Recruitment and Treatment

The study was approved by The Columbia University Institutional Review Board and is HIPPA compliant. Written consent was obtained by each subject. Women over the age of 18 who were diagnosed with stage II or stage III breast cancer and were to undergo NACT were eligible for the study. The primary tumor had to be at least 1 cm in diameter and women with metastases were not eligible. Each patient receives twelve weekly treatments of Paclitaxel followed by four cycles of doxorubicin and cyclophosphamide (AC) given every other week. There are six DOT imaging time points: baseline, cycle 3 and 5 of Paclitaxel, before AC, cycle 1 and 2 of AC, and before surgery (Fig 2.1).

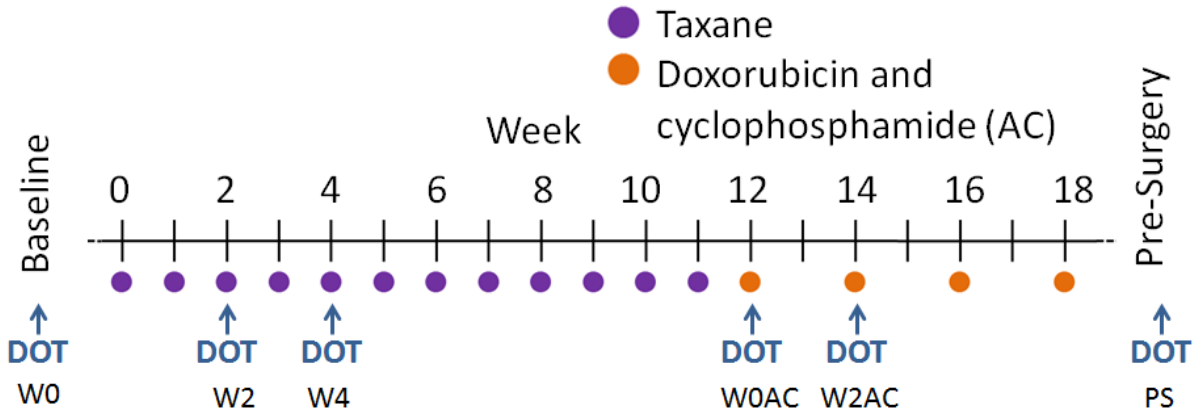


Figure 2.1 Timeline for subject treatment and imaging.

These imaging time points were chosen for two reasons. First, we wanted to study tumor progression as a patient received chemotherapy throughout the entire course of treatment. Second, we wanted to observe if there are time points during chemotherapy that can be used to predict the overall pathological response of the subject. Imaging time points were selected close to the beginning of treatment and at points when the chemotherapy agents were changed. This was designed so that we could observe the

beginning of each treatment to analyze early effects. Response to treatment was based on histology performed after surgery.

2.2.2 Baseline Biomarker

Before the beginning of treatment, diagnostic biopsies for each patient were tested for specific biomarkers that may assist in determining optimal treatment. Current clinical care for breast cancer patients include determining biomarkers, such as estrogen receptor (ER), progesterone receptors (PR), and amplification of HER-2/*neu* (HER2),as well as, a widely used test of Ki67 staining.

ER and PR are hormone receptors and are used to predict the response to anti-estrogen therapy.[61] Specifically, ER-α, an isoform of ER, is responsible for mammary gland development and tumorigenesis.[62] Positive ER and PR statuses have a fairly good clinical prognosis, while negative ER and PR status have a poor prognosis. HER2 is an oncogene, which is linked to negative ER/PR status. HER2 is also part of the standard work-up on newly diagnosed patients. Amplification of HER2 is associated with a poor prognosis.[61] Another biomarker we observe for this study is Ki67, which is a measure of cell proliferation in the cancer. This is a widely used biomarker that has been used to predict response to chemotherapy. [61]

There has been sufficient evidence to support that the pathologic response predicts improved clinical outcome in specific BC subtypes, such as triple negative BCs (negative for ER, PR, and HER2).[63] Therefore, combining these biomarkers with optical imaging features may increase the accuracy of our prediction models to help determine the pathological response to treatment before surgery.

2.2.3 Clinical Exam Measurements

Before each cycle of chemotherapy, patients would receive clinical exams and have their tumors measured extrinsically, by use of calibers. The maximum dimension of the tumor was used for analysis. The clinical measurements were taken to give the physicians an indication of how well the patient was doing without using any imaging or biopsy techniques. We used this information to compare to our DOT measurements to conclude if DOT can perform better than these standardized clinical measurements.

2.2.4 Residual Cancer Burden

At the end of the treatment breast cancer surgery was performed and any remaining tumor mass was removed and evaluated by histopathology. For each person the residual cancer burden (RCB) score was calculated. The RCB score takes into account the size and proportion of the residual primary tumor and number and diameter of residual lymph nodes.[63]

The following formula is used to calculate the RCB score:

$$RCB = 1.4\left(f_{inv} d_{prim}\right)^{0.17} + [4(1 - 0.75^{LN}) d_{met}]^{0.17} \quad (2.1)$$

where f_{inv} is the proportion of the primary tumor bed that contains invasive carcinoma, LN is the number of axillary lymph nodes that contain cancer, d are the dimensional measurements of the primary (d_{prim}) or metastatic cancer in the lymph node (d_{met}). The primary tumor dimension was determined by using the two largest dimension and calculating: $d_{prim} = \sqrt{d_1 d_2}$.

The RCB score is a score given on a continuous scale, but then can be further binned into classes. A person was considered a pathologically complete responder (pCR) if the RCB ≤ 0.5 (RCB-0). A patient with RCB > 3.28 (RCB-III) was considered a non-

responder. For values in between patients were classified as partial responders (RCB-I for 0.5< RCB≤1.36, and RCB-II for 1.36< RCB≤3.28).[64]

Both pCR and RCB have been identified as surrogate endpoints for clinical outcome, as pathologic response (i.e. pCR or low RCB) associate with a favorable prognosis including longer disease free survival and overall survival.[24, 63-65] Yet, pathologic response rates are low. In a pooled analysis of over 6,000 patients administered neoadjuvant anthracycline-taxanes, pCR was achieved in only 15% in all BC subtypes. “Responders” as determined by a RCB score of 0 or 1, occur in approximately 40% of patients with BC after NACT, with “non-responders” (RCB = 2 or 3) identified in about 60%. RCB independently predicts for distant BC relapse-free-survival in multivariate models that include pCR.[64] Recent NACT guidelines recommend routinely checking RCB.[65]

For the NACT monitoring study, we used the RCB score and RCB class as an endpoint to compare with our DOT features. Our goal was to correlate the RCB score with our DOT features and to also predict pCR.

2.3 Overview of DOT Imaging

2.3.1 Instrumentation and Patient Interface

To image subjects we used a custom built digital dynamic DOT system. Our current breast DOT system is a continuous wave dynamic imaging system that emits four wavelengths. Two of the wavelengths (765 and 808 nm) are modulated at 5 kHz, and the other wavelengths (827 and 905 nm) are modulated at 7 kHz. Modulated light sources were used to implement a lock-in detection system to reduce noise.[22, 66]

Due to the large geometries that are being imaged, the signals received by the detection unit vary greatly. To combat this problem, the imaging system was designed to have a large dynamic range (~120 dB). Each source detector pair is allocated a gain setting that gives the acquired signal the optimal signal to noise ratio (SNR) without saturation. The gain settings determine how much amplification is necessary for the signal from each source-detector pair.

Table 2.1 Summary of DOT breast imager parameters.[22]

Parameter	Value
Sources	32 per breast, 64 total
Detectors	64 per breast, 128 total
Wavelengths	~5 mW at 765 nm, 808 nm,827 nm, 905 nm
Temporal Response	1.7 Hz – 111 Hz
Dynamic Range	~ 10^8 (158 dB)
System Size	104 cm x 79 cm x 66 cm
Power Consumption	3.5 Amps @ 120 V AC

There are a total of 64 sources and 128 detectors (see Table 2.1). The patient interface is comprised of two sets of four concentric rings that hold the source and detector fibers in place, and can be adjusted to accommodate different breast sizes. Each set of the rings holds a total of 64 fibers in which all fibers are detectors, but half are collocated sources. Fibers are placed in a source-detector-source-detector pattern around each ring. There are four rings that vary in size with a diameter of 4, 8, 12, and 16 cm and hold 8, 12, 16, and 28 fibers, respectively. The angle between the fibers also changes as the size of the rings increase (60˚, 45˚, 30˚, 15˚) to assure fibers are normal to the tissue and maximum coverage of the breast is assured. Larger rings can be removed to accommodate smaller breast sizes. However, frame rate is dependent on the number of sources and detectors used to image, so if a ring is removed the system images at a

faster frame rate. We preform 3D imaging of both breasts simultaneously with a frame rate of 1.7 Hz when all sources and detectors are used.[22] The fast acquisition time allows us to view hemodynamic changes that occur due to stimuli such as a breath hold (~30 sec).

Figure 2.2 Image of the DOT breast imager with translating ring interface.

Our DOT system is capable of dynamic imaging of both breast simultaneously, which allows us to compare the healthy and tumor bearing breasts. There is no compression or extrinsic contrast agents necessary, which allows for more comfortable imaging experience for the patient. Also, the system has a fairly quick acquisition time which allows us to acquire about 1000 frames within 8-10 minutes.

2.3.2 Patient Imaging

The height of each ring was adjusted for each subject to assure that all of the optical fibers made contact with the tissue. The subject placed her breasts inside the translating rings and the gain settings were automatically determined for each source-detector pair through custom software written in LabView. The gain settings determine

how much amplification is necessary for the signal from each source-detector pair. After optimizing the gain settings for the subject, imaging began with a baseline measurement of about 2 minutes in which the patient remained stationary. Next, the patient was instructed to hold their breath for thirty seconds, and then was given about 90 seconds of recovery time. Each patient performed two to three breath holds, and the total imaging time was around 8 minutes (Figure 4). Once the subject imaging was done and she was removed from the imaging probe, a reference solution was imaged for about a minute. The reference solution was composed of 20 % Intralipid and 1% India ink, which was diluted to have similar optical properties as breast tissue.

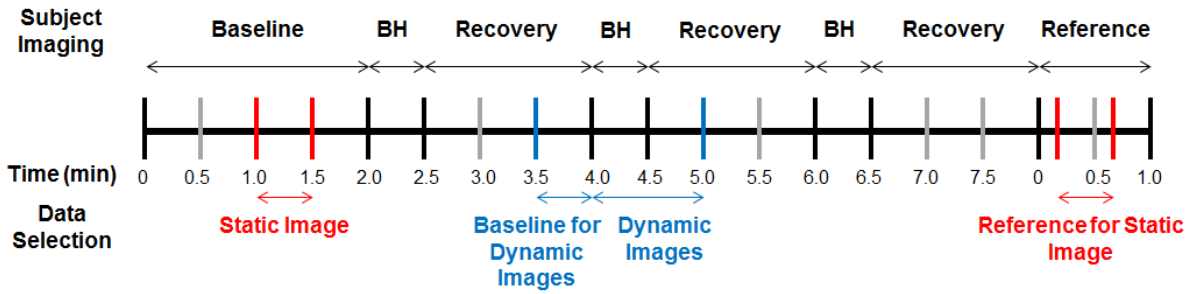


Figure 2.3 Timeline for subject imaging and an example of how data is selected to reconstruct images. Baseline refers to the period of time that the subject remains stationary. Subjects perform three breath holds (BH) for about 30 seconds and given 90 seconds to recover. Reference solution is imaged for one minute after the subject imaging is over.

The acquired data was reconstructed to create both static and dynamic images. Static images were reconstructed from the measurements acquired during the subject's baseline imaging and the reference solution imaging. By using a reference solution of known optical properties, static 3D images of oxy-hemoglobin concentration, deoxy-hemoglobin concentration, and water fraction were made. To create the dynamic images, one of the breath holds was selected for reconstruction. The baseline measurement for the dynamic images were taken 30 seconds before the patient began holding their breath

and was used to calculate percent change of oxy-hemoglobin and deoxy-hemoglobin concentrations for the duration of the breath hold, as well as, 30 seconds of the recovery time. About one dynamic image per second was reconstructed so that the hemodynamic effect of the subject's breath hold could be analyzed.

Other studies have used the Valsalva maneuver, mechanical compression, or contrast agent, such as ICG to observe hemodynamic effects of breast tumors[67, 68]. However, these pose a couple of challenges. One group needed to exclude a third of their subject population due to motion artifact caused by the Valsalva maneuver[68]. This maneuver requires to breathe out for a certain length of time which could cause motion issues during imaging. However, with the breath hold, the subject is not constantly moving during the imaging. Additionally, there is no extrinsic contrast agent that could cause some discomfort to the patient.

2.3.3 Mesh Construction

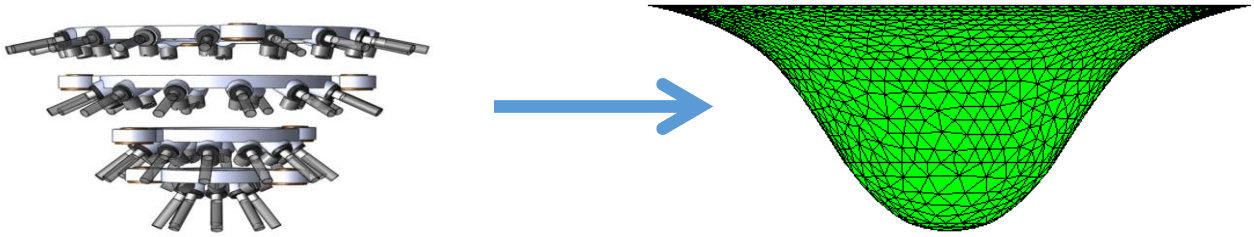


Figure 2.4 Computer rendering of translating ring interface. The location of each source and detector is known, as well as, the distance between rings. This information can be used directly to create 3-dimensional meshes for reconstructions.

Meshes for image reconstructions were created using the geometry of the rings and known locations of each source and detector (Fig 2.4). Therefore, meshes were generated custom made for each individual subject, which improves the accuracy of the reconstructed image. The meshes extend 2 cm past the largest ring to account for the

effects that the chest wall might have on the reconstruction. Meshes contained approximately 60,000 to 70,000 elements.

Meshes are 3-dimensional volumes that are segmented into voxels. Each of the voxels are an unknown to which the reconstruction algorithm assigned a value (i.e. μ_a, μ_s, etc.). The meshes are also used to predict how light travels through the medium that has been assigned specific optical parameters. Mesh generation is an essential step in reconstructing optical images, since the distance between the source and detector are one of the key factors for predicting how light travels through the medium. Additionally, the meshes are used to compile the final 3D DOT image.

2.3.4 Reconstructions

The imaging data was pre-processed and reconstructed to yield 3-dimensional images and movies of the whole breast. A PDE-constrained multispectral imaging method that uses the diffusion approximation as a model for light propagation.[69] Static and dynamic reconstructions were performed on previously developed codes within the lab. More detail about the specifics of the reconstruction algorithms can be found in chapters 3 and 4. Otherwise, optimization schemes similar to the reconstruction algorithm structure explained in Chapter 1 were utilized.

2.4 Statistics

2.4.1 Standard Statistics

The mean and standard deviation were determined for each of the previously described features. One way ANOVAs were performed for comparison of more than two groups. Otherwise, two-way unpaired student t-tests between designated groups was

performed. Linear regressions of the RCB score versus different DOT features in which the correlation coefficient and p-value were determined.

2.4.2 Sensitivity Analysis

Classifier performance can be determined with the help of a confusion matrix. A confusion matrix displays how many of a given population was predicted correctly for a particular feature being observed (i.e. how many women were correctly diagnosed with breast cancer using x-ray mammography). Figure 2.5 shows how the confusion matrix is constructed.

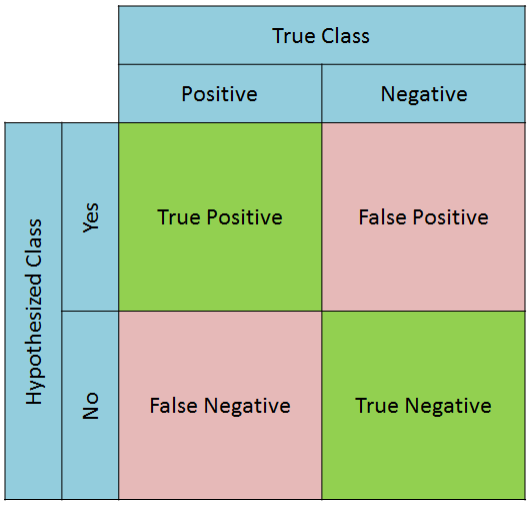


Figure 2.5 A template for a confusion matrix.[70]

Subjects are classified based on an experimental feature (hypothesized class), but are compared to a gold standard of diagnoses (True Class). Using our previous example, a true positive would be a subject that had breast cancer and was classified as having breast cancer. A true negative would be a subject that did not have breast cancer and was correctly classified as have no breast cancer. A false positive (FP) or false negative (FN) were subjects that were not classified correctly. The sensitivity and specificity of the confusion matrix are determined using:

$$Sensitivity = \frac{TP}{P_D} \tag{2.2}$$

$$Specificity = \frac{TN}{N_D} \tag{2.3}$$

where TP and TN are the number of subjects that were true positive and true negative, respectively. P_D were the number of positively diagnosed subjects and N_D were the number of negatively diagnosed subjects. The accuracy is determined by equation 2.4.

$$Accuracy = \frac{TP+TN}{N_D+P_D} \tag{2.4}$$

Additionally the positive predictive value (PPV) and negative predictive value (NPV) can be calculated using:

$$PPV = \frac{TP}{P_T} \tag{2.5}$$

$$NPV = \frac{TN}{N_T} \tag{2.6}$$

where P_T is the number of subjects that were classified positive and N_T were the number of subjects tested negative.

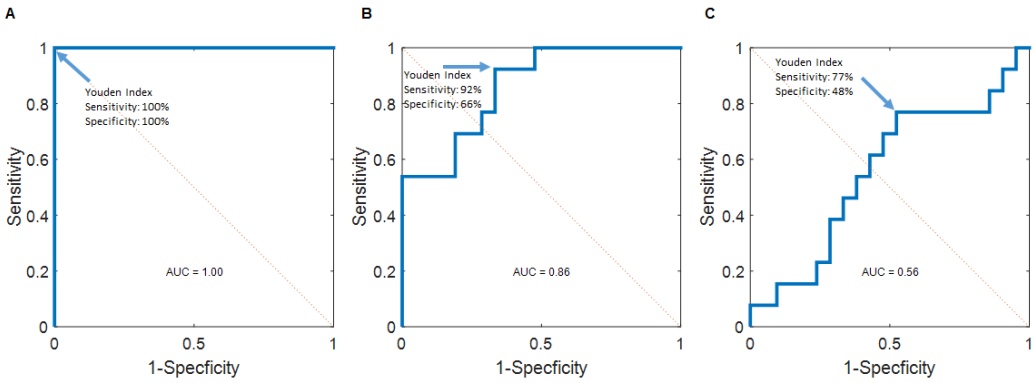


Figure 2.6 Examples of ROC curves with the sensitivity and specificity at the Youden index. An ideal ROC curve with an area under the curve (AUC) of 1 (A), a standard ROC curve with an AUC of 0.85, and a very poor ROC curve with an AUC of 0.56 (C).

Receiver operating curves (ROC) compare the specificity and sensitivity of a diagnosis test and are efficient at determining the efficiency of classification techniques.

The sensitivity and specificity of several threshold values within a given feature are tested and the threshold that has the maximum value to the equation 1-(Sensitivity + Specificity) is considered the Youden index, which is the optimal cutoff point for the classifier. The area under the curve of the ROC can be determined. An AUC of 1 is ideal for a classifier (Fig 2.6A). Examples of ROC can be seen in Figure 2.6 in which there are different AUCs and the location of the Youden index are identified with the corresponding sensitivity and specificity.

Thus, the sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) were determined from the Youden index of the ROC curve. The sensitivity within this thesis denotes the percentage of the more responsive group that was diagnosed correctly. Conversely, the specificity denotes the percentage of subjects in the least responsive group that were categorized correctly.

Univariate and multivariate discriminant analysis with cross-validation was also employed and described in Chapter 7.

2.5 Summary of Subject Population

Table 2.2 Subject summary and pathologies.

	Avg ± Stdev			**Range**
Age (Years)	49 ± 12			(29-73)
BMI	29.8 ± 6.0			(18.3-44)
Tumor Size (cm)	4.4 ± 2.7			(1.1-12)
Menopausal State	Pre-Menopausal			Post-Menopausal
Number of Subjects	20			16
Cancer Type	Invasive Ductal		Invasive Lobular	Mixed Ductal & Lobular
Number of Subjects	32		4	1
RCB Score	RCB 0	RCB I	RCB II	RCB III
Number of Subjects (Percent of subject population)	14 (38.9%)	2 (5.5%)	14 (38.9%)	6 (16.7%)

There were forty women recruited and imaged for the study with four subjects being excluded due to drop outs or machine malfunctions. Four patients were excluded. One patient completely stopped NACT after the first imaging. For 2 subjects, instrumentation failure during the imaging session resulted in corrupted data sets that could not be used. Finally, one subject opted to receive radiation therapy in addition to the traditional NACT therapy and therefore was removed from the analysis.

Table 2.3 Number of subjects in each response category for each week.

# of Subjects	*Baseline*	*Week 2*	*Week 4*	*Week 0AC*	*Week 2AC*	*Pre-surgical*
RCB 0	14	13	12	10	11	7
RCB I	2	2	2	2	2	2
RCB II	14	13	14	13	12	11
RCB III	6	6	6	6	6	6
Total	36	34	34	31	31	26

* Subject imaged during week 3 instead of week 2 (1 RCB 0 subject and 2 RCB II subjects)
** Two subjects imaged at third AC cycle instead of the second cycle.

As seen in Table 1, the average age of the women was 49 ± 12 in which the youngest woman was 29 and the oldest was 73. The average body mass index (BMI) was 29.8 ± 6.0. There were 20 pre-menopausal subjects and 16 post-menopausal women. Most women had invasive ductal carcinoma (32 subjects). There were four women with invasive lobular carcinoma and one subject with mixed ductal and lobular. In the discrete scale of RCB score, 14 subjects had an RCB of 0 (pCR), 2 subjects had an RCB of I, 14 subjects had an RCB score of II, and 6 subjects had an RCB of III. Not all subjects were imaged at every time point. Table 2 shows a summary of the number of subjects that were included in the analysis for each week for each response group.

2.6 Novelty and Uniqueness of Current Study

The current study has several features that make it unique. First, our imaging system solely operates as an optical system. There have been studies of dual-imaging systems that combine near-infrared imaging with US, MG, and MRI.[25, 26, 58, 71-75] The advantage is that commonly used imaging modalities that physicians are comfortable using are being used the combined information that near-infrared imaging has to offer. However, when it comes to treatment monitoring these imaging modalities lack the versatility or efficacy that is needed to image patients multiple times safely at a reasonable cost. DOT uses non-ionizing light sources that allow for multiple imaging session within a short period of time. Therefore, unlike x-ray imaging, frequent DOT measurements can be performed without adverse effect. DOT is also sensitive to parameters that change over the course of treatment, such as tissue vascularity or hypoxia.[3, 48]

There are several facets that make our current study unique to other imaging studies that have been published.[26, 57, 59, 72-81] First, all subject receive the same chemotherapy regimen. Other chemotherapy studies have subjects on different chemotherapy regimens, which could lead to different responses and thus make predicting the response to certain treatments difficult. Second, we have six imaging time points spaced throughout the treatment with imaging concentrated towards the beginning of treatment and after switching chemotherapy agents. We have recruited 40 subjects and analyzed 36 subjects, which is one of the larger optical studies performed to date.

Most importantly we are collecting both static and dynamic information which could lead to a more complete picture of what the tumor progression is under the treatment circumstances. We can possibly use these two pieces of information to accurately monitor the tumor or possibly predict the response early within the treatment.

2.7 Conclusion

We have designed a study that allows for the collection and analysis of multiple facets of tumor progression throughout NACT. Through the use of a versatile dynamic DOT imaging system, we have collected static and dynamic optical information. For the remainder of this thesis, I present how this information was processed and analyzed, as well as, the overall results we have found from this longitudinal pilot study.

Chapter 3: Static Data Processing

3.1 Introduction

Since DOT is sensitive to hemoglobin concentrations, it is capable of characterizing changes in vascular structures and blood perfusion in tissue. Several studies have shown that there is a change in optical properties over time during the course of NACT. [26, 48, 57-60, 72-81] We have used a DOT breast imaging system previously developed in our group that is capable of acquiring 3D transmission data from both breasts simultaneously to measure spatially dependent concentrations of [HbO_2], [Hb], [HbT], and water fraction (WF), and oxygen saturation (SO_2%) on subjects receiving NACT. We plan on looking at the percent change from the baseline of these parameters over time to track tumor progression and to correlate these changes with the pathological response of the subject.

3.2 Challenges to overcome

Since the DDOT system was set up for dynamic imaging, our previous studies had focused on the hemodynamics of breast cancer. However, this has been the first study in which we planned to look at static imaging with the aid of a reference phantom. Therefore, several adjustments had to be made and techniques developed for the current data set to obtain accurate static images.

The original code being utilized was a PDE-constrained multispectral imaging method that uses the diffusion approximation as a model for light propagation. However, this code would work very well for some patients and pose huge problems for other patients. There would be large artifacts, false positives in the contralateral breast, and

indeterminate cancer locations in our original protocol setup. Also, the tumor-bearing masses would display blobs in the middle of the images which would be caused the by regularization within the reconstruction algorithm. Regularization was kept constant for all patients but would have needed to be changed base on mesh size. There is no method of determining this efficiently.

Once we switched to the DCT reconstruction method, the artifacts at the surface of the images became clearer. As a result, there needed some method of artifact removal incorporated into the pre-processing stage. We also added a method for determining the initial conditions for each reconstruction. Last, there was no protocol put in place for the quantification of the images, so a new method had to be developed.

3.3 Pre-processing

3.3.1 Overview

Figure 3.1 shows an overview of the pre-processing protocol. First, 30 seconds of the raw data are averaged for the patient and reference measurement and saved to text files. There is no additional pre-processing methods performed on the raw data.

Next the forward model was run and the input and setup files were created. The input files are the calculated data using the following equation:

$$z_{tar}^{\lambda} = \frac{M_{tar}^{\lambda}}{M_{ref}^{\lambda}} P_{ref}^{\lambda} \tag{3.1}$$

where M_{tar}^{λ} is the subject measurement for a specific wavelength, M_{ref}^{λ} is the reference measurement for a specific wavelength, P_{ref}^{λ} is the prediction measurement for a specific wavelength from the forward model, and z_{tar}^{λ} is the input value for the reconstruction. The

setup files contain information for the reconstruction including the initial condition. The evolutionary strategy was run and used to change the initial condition in the setup files.

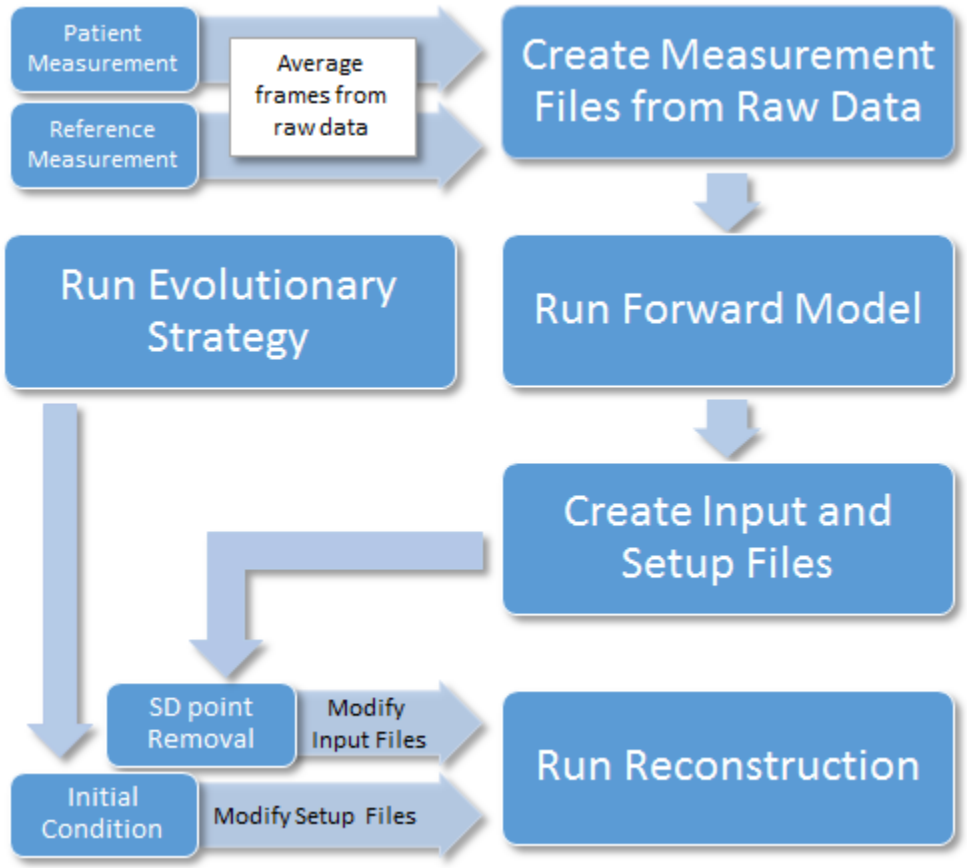


Figure 3.1 Overview of pre-processing steps to run static reconstructions.

Also, the source-detector (SD) points are analyzed, the poor points are removed, and the input files are modify to reflect the changes. Last, the reconstruction is run. All the time points and the left and right breasts are pre-processed and run together. This saves time and allows all the time point images to be processed the same way.

3.3.2 Evolutionary Strategy

Before utilizing evolution strategies for the static reconstructions all patients were given the same initial value for the reconstruction algorithm. There has been evidence to support that the initial value can have a huge impact on the outcome of the inverse

problem.[30] Additionally, the optical properties of breast tissue can vary drastically according to age or BMI.[82] Therefore, there was a need for a more individualistic means of inputting the initial conditions into the reconstruction algorithms. One simple solution was to use pre-written evolutionary strategy algorithms to determine an estimate of the breast tissue properties.

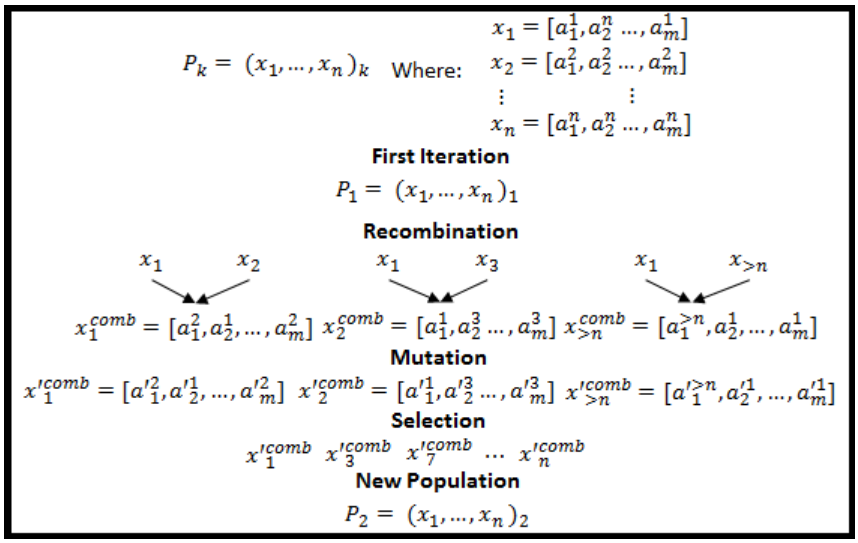


Figure 3.2 Evolution Strategy uses an iterative process of recombination, mutation, and selection of a set of individuals P_k to determine optical properties.

The initial guess for the full tomographic images is predetermined beforehand using evolutionary strategies. Briefly, evolution strategies uses principles of nature such as survival of the fittest and heredity to determine the optimal solution for an ill-posed problem. This is an iterative process that starts with a parent population of $P_k = (x_1, \dots, x_n)_k$ where k is the number of iterations, P is the parent population, x_i is one of the individuals of the population and n is the total number of individuals (Figure 3.2). Each individual is represented by $x_i = [a_1, \dots, a_m]$ where a_j is one of the optical properties that needs to be solved out of *m* optical properties.

The properties of the individual are then recombined with another individual in the parent population. This then becomes the beginning of the child population that contains $C_k = [x_1^{comb} x_2^{comb}, \ldots x_{>n}^{comb}]$ Afterward a mutation occurs to each property for each of the individuals in the child population where $a_j^{mut} = a_j' + \sigma_j \cdot N_j(0,1)$. Sigma represents the mutation step size of the jth individual and $N_j(0,1)$ is a random number from a normalized distribution. After the mutations are completed the forward model is run and the outcome is compared to the measurement data taken in the experiment. The objective function for each child is found and the n lowest objective values are chosen for the new population for the next iteration. This is repeated until one individual reaches a threshold for the objective function value.[83]

The forward model for the ES is based on reflectance measurements and is dependent on the absorption coefficient (μ_a), the reduced scattering coefficient (μ_s'), and source-detector distance (d).

$$R(d)_{dc} = \frac{1}{4\pi\mu_t'}\left[\frac{\left(\mu_{eff}+\frac{1}{r_1^2}\right)\exp(-\mu_{eff}r_1)}{r_1^2} + \frac{\left(\frac{4}{3}A+1\right)\left(\mu_{eff}+\frac{1}{r_2}\right)\exp(-\mu_{eff}r_2)}{r_2^2}\right] \quad (3.2)$$

where

$$r_1 = \sqrt{\left(\frac{1}{\mu_t'}\right)^2 + d^2} \quad (3.3) \qquad r_2 = \sqrt{\left(\frac{\frac{3}{4}A+1}{\mu_t'}\right)^2 + d^2} \quad (3.4)$$

$$\mu_{eff} = \sqrt{3\mu_a\mu_s'} \quad (3.5) \qquad \mu_t' = \mu_a + \mu_s' \quad (3.6)$$

where μ_{eff} is the attenuation coefficient, μ_t' is the transport coefficient and A is the internal reflection parameter that takes into account the reflective index mismatch at the air-tissue interface.[83]

To determine the initial condition of the static reconstruction only a fraction of the subject data was used in the evolutionary strategy. The patient interface on the DDOT system has four concentric rings. Only the data acquired from the second ring of the patient interface was used for the ES method. The second ring was used since this was the ring that was most likely to have the least trouble making complete contact with the subject's breast and therefore, would have the least artifacts so that the ES method would not be influenced by bad data. In addition, this increases the speed of the ES code by reconstructing a uniform 2D area instead of a 3D volume.

3.3.3 Data Point Selection

The data point selection method had a number of criteria that needed to be considered such as:

1. Keeping the same number of data points for the tumor-bearing and healthy breast.
2. Keeping the same number of data points for all time points in the longitudinal study.
3. Identifying and removing irrelevant data points.
4. Automate the removal process.

In many cases, there are sources or detector fibers that were either not calibrated properly, not making contact with the patient's skin, or just not working properly. Yet, if we look at the target measurement divided by the reference measurement, such as in Figure 3.3, the poor data points can be identified. In Figure 3.3A, there is an example of the calculated data for all source-detector pairings for a single wavelength. The data seems to have not observable issues with data, and most likely would not have issue during reconstruction. However, in Figure 3.3B, there is a whole source and detector that seems to have very little signal. The aberrant data points would cause a large problem in the reconstruction as seen in Figure 3.4A.

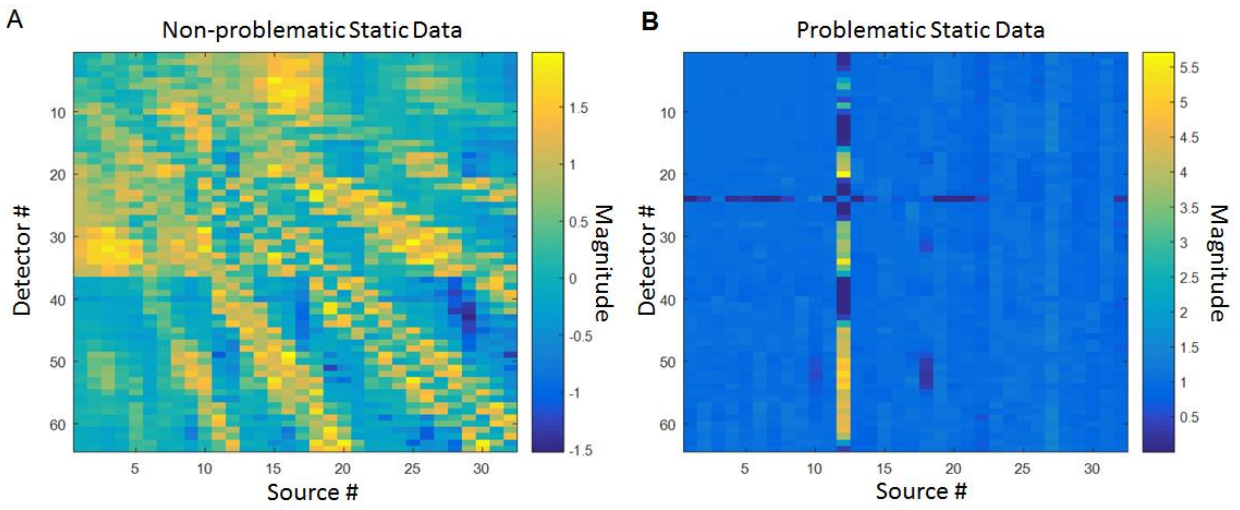


Figure 3.3 An example of a non-problematic static data set (left) and a problematic static data set (right). The magnitude is $\log_{10}(M^{\lambda}_{tar}/M^{\lambda}_{ref})$. The problematic data set needs source 12 and detector 24 need to be removed from the reconstruction input.

Figure 3.4 shows the example of a phantom measurement. During the phantom measurement, source 12 (detector 24) was covered with black material. The source concurrently acts as a detector. In Figure 3.4A, there is an artifact in the location of the source-detector location. After the data points that are recorded when the source was illuminated and the data points of the corresponding detector were removed from the reconstruction input, the artifact disappeared from the image (Fig. 3.4B).

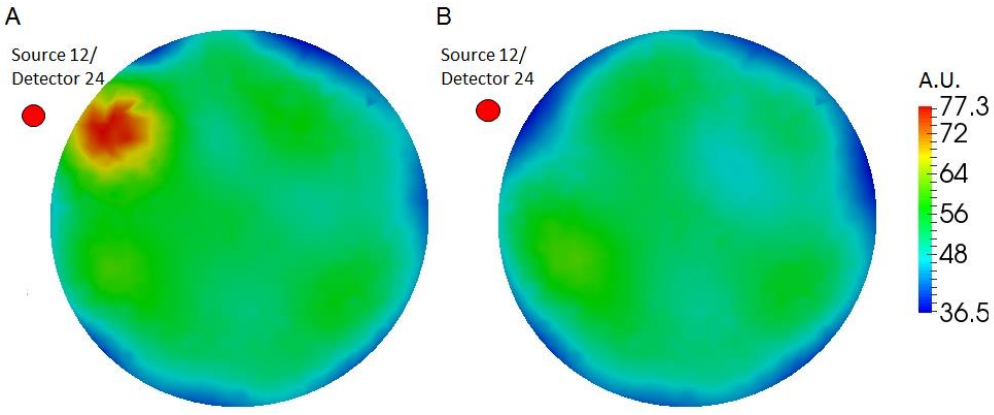


Figure 3.4 Coronal slices of the same phantom measurement before (A) and after (B) source 12 (detector 24) had been removed from the input to the reconstruction algorithm.

Therefore, we developed a series of criteria to remove poor data points. The magnitude of the point was calculated by using $\log_{10}(M_{tar}^{\lambda}/M_{ref}^{\lambda})$ for each source-detector pair and was examined by the criteria. A data point was removed from the reconstruction if they fit under the following criteria:

1. Source-detector points that were greater than $Threshold_p^1 = P_{avg} + \sigma * P_{StDev}$ or less than $Threshold_p^2 = P_{avg} - \sigma * P_{StDev}$, where σ is a constant for the degree above and below the mean (P_{avg}) based on the standard deviation (P_{StDev}).
2. The median magnitude of the source across all detectors was calculated (S1, S2, S3…S32). All data points for a particular source that had a median greater than $Threshold_S^3 = S_{median} + \beta * IQR_{Source}$ or less than $Threshold_S^4 = S_{median} - \beta * IQR_{Source}$, where β is a constant for the degree above and below the median base on the interquartile range (IQR) and S_{median} was the median of set (S1,S2, S3,…S32).
3. The median magnitude of the source across all detectors was calculated (D1, D2, D3…D32). All data points for a particular detector that had a median greater than $Threshold_D^1 = D_{median} + \beta * IQR_{Detector}$ or less than $Threshold_D^2 = D_{median} - \beta * IQR_{Detector}$, where β is a constant for the degree above and below the median base on the interquartile range (IQR) and D_{median} was the median of set (D1,D2, D3,…D32).

If more than a third of the detectors for a single source meet criteria 1 then the whole source is removed from the reconstruction for all wavelengths. If more than a third of the sources for a single detector meet criteria 1 then the whole source is removed from the reconstruction for all wavelengths. The corresponding detector to a removed source is

removed. If a removed detector had a corresponding source, the source is removed. If the difference in the percentage of data points removed was greater than 7% for any two imaging time points, then σ was decreased by 0.1 until the data removal fit the condition or the lowest percent difference was found. The algorithm was applied to all the images of a single patient to assure that they were pre-processed in the same way. Afterwards the modified input files were used for the reconstruction.

The mean and standard deviation were considered for data points because using the above thresholds would remove any outliers or near outliers from the average of the data (criteria 1). The median of each source and detector was found to avoid influences of single data points that would be removed by criteria 1. Then the outliers of these medians were used to remove just sources or detectors that were vastly different from the rest of the sources or detectors (criteria 2&3).

The percent difference was chosen around 7% since there needed to be some flexibility on the amount of data used per image, but this percent difference still need to be low. The overall aim was to have a percent difference less than 5%, but for the algorithm this was increased slightly since 5% was sometimes difficult to achieve. The percent difference was due to the different source-detector points that would be removed from the different time points since different sources would pose problems at different time points. Therefore, the same amount of data removal from each image was extremely difficult to achieve. If there the same points were removed from each image, the majority of the data would be removed and could not be used for reconstruction.

The constants σ and β were tunable before starting the algorithm. The tunable parameters were necessary for two reasons. One, not all subjects, would reach the

percent difference between two images threshold and the initial parameters would need to be changed. Also, if the parameters were too conservative (i.e. $\sigma = 2$ and $\beta = 1.5$) the majority of the data could be removed from the reconstruction, which would yield poor results. Therefore, the recommended values for the parameters should be $\sigma > 2$ and $\beta > 1.5$. Otherwise, the algorithm would start removing relevant data that statistically would not count as an outlier.

3.4 Discrete Cosine Transform Reconstructions

The traditional method for optical image reconstructions are element based, which means the number of unknowns in the problem are the number of elements in the mesh. However, there are usually less measurements available then unknowns in practical applications causing a highly ill-posed problem. The element based reconstruction methods are then sensitive to measurement noise, the initial guess, and the regularization parameter. The regularization parameter would have to be optimized for each mesh and since the current experiment involves multiple subjects each with different size meshes, an efficient method for determine the regularization parameter for multiple meshes was not possible. As seen in Figure 3.5B, the regularization parameter also caused a blob formation within the center of the reconstructed image, which made tumor positioning difficult to determine. Therefore, we changed over to a DCT Reconstruction method for our static breast reconstructions.

The reduced space basis function neural network method for the DOT image reconstruction was previously developed in our lab. Briefly, this method uses a basis function to reduce the solution space, so instead of solving each element which can range from the 8k to10k elements, the algorithm now solves the basis function, which reduces

the unknowns to about 1000. The basis function of discrete cosine transformation (DCT) was used.[84]

3.5 Static Image Improvement

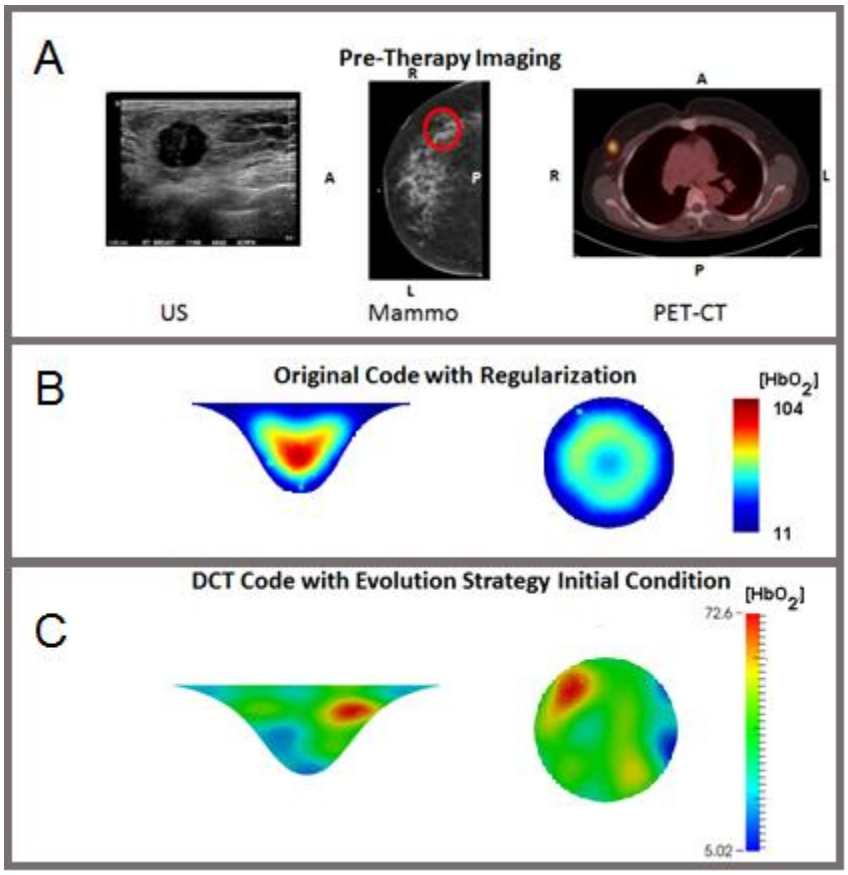


Figure 3.5 Comparisons of traditional imaging methods, DOT images, and improvements made to the reconstruction scheme. A 47 year old pre-menopausal women who was diagnosed with poorly differentiated ductal carcinoma with a tumor of about 2 cm located in the 11 o'clock position 6 cm behind the nipple. (A) Images show pre-therapy images taken from different modalities. (B) Sagittal (3 cm from center) and coronal slices (6 cm from nipple) are shown using the original code. (C) Same sagittal and coronal slices as above, but reconstructed using DCT code with initial condition. All images are baseline imaging for the tumor-bearing breast (with the exception of PET).

We now use a DCT reconstruction algorithm that solves for a set number of DCT components instead of solving for each individual mesh element. This has eliminated the need for modifying the regularization parameter, and has increased the speed of reconstruction. We also started to use an evolution strategy algorithm to determine the initial condition parameters of the reconstruction. Figure 3.5 shows the how the baseline imaging for a single patient has changed from the original to the current protocol.

In the pre-therapy images (fig. 3.5A), the tumor is visible in all three images modalities: US, MG, and PET-CT. The tumor was located in the 11 o'clock position in the right breast 6 cm from the nipple. In the original DOT images (fig 3.5B), the tumor appears in the center and does not correlate with the other imaging. Yet, in Figure 3.5C, we can see the tumor in the 11 o'clock position about 6 cm from the nipple. Therefore, the addition of these pre-processing tools has assisted in improving the DOT static images.

3.6 Post-processing

3.6.1 Volume of Interest Selection

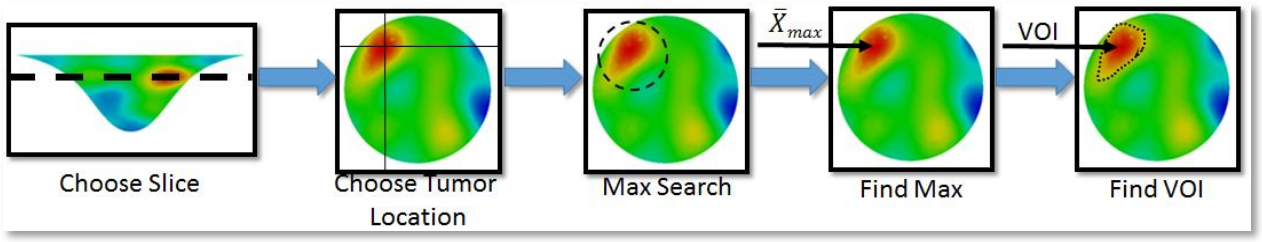


Figure 3.6 Flow diagram of quantification of static images. The VOI are all the voxels that have a value of M-SD < V_i< M.

After the reconstructions were performed the images were analyzed. To quantify the images, a semi-automated custom written program was used to locate the tumor. A user would look at DOT images and asked to locate the tumor (Figure 3.6). In order to

reduce bias, the algorithm then looked for the maximum value (M) within a 2cm radius. The location of the maximum location was $\bar{X}_{max} = (x_{max}, y_{max}, z_{max})$. Additionally, the standard deviation (SD) of the volume was calculated. The volume of interest then became all voxels with value V_i within the region that satisfied the condition M-SD ≤ V_i ≤ M. Then the weighted average of the volume was taken. To calculate the healthy breast the weighted average of the whole volume was used. The normalized values were the tumor region divided by the healthy breast values.

3.6.2 Features

The reconstructions solve for oxy-hemoglobin ([HbO2]), deoxy-hemoglobin ([Hb]), and water fraction (WF). Total hemoglobin ([HbT]) was calculated summing [HbO2] and [Hb] ($[HbT] = [HbO_2] + [Hb]$). Oxygen saturation (SO_2%) was calculated using:

$$SO_2\% = 100 * \frac{[HbO_2]}{[HbO_2]+[Hb]} \quad (3.7)$$

The normalized values were calculated in two ways. First, the parameters were divided by the parameter values for the healthy breast ($[HbO_2]_{T/H}$, $[Hb]_{T/H}$, etc.). Second the parameters were divided by the weighted average for the non-tumor region in the tumor bearing breast ($[HbO_2]_{T/nT}$, $[Hb]_{T/nT}$, etc.). Additionally, the percent change was calculated from the baseline imaging for both the original optical features and the normalized values (i.e. %Δ[HbO2], %Δ$[Hb]_{T/H}$, etc.).

3.7 Conclusion

By utilizing the DCT code and the source-detector removal conditions the quality and accuracy of the image was greatly improved. The automated pre-processing and batching all of a single patients images saves time and allows for consistency.

Additionally, we have established a method to extract the tumor region without interference from artifacts that also reduces the bias of the user. Then we analyzed how the above mention features correlated to the pathological response of the subject and how they changed over time.

Chapter 4: Static DOT Results

4.1 Introduction

Diffuse optical tomography (DOT) can create 3D maps of various tissue parameters using transmission measurements with non-ionizing, near-infrared light. From these measurements, endogenous chromophores such as oxyhemoglobin, deoxyhemoglobin, water, and lipid can be used to differentiate malignant from normal tissue.[25, 37-39]

Due to the sensitivity of optical imaging, multiple groups have employed this technology for monitoring and predicting response to NACT. Ueda et. al. correlated the pCR rates to baseline tumor oxygen saturation measurements by using a handheld spectroscopic measurement system.[85] A subset of these patients received NACT combined with bevacizumab or trastuzumab. They observed an elevated baseline oxygen saturation (SO_2%) level corresponded highly with a pCR for treatment. Zhu et. al. observed 32 subjects undergoing NACT using a combined ultrasound DOT imaging.[58, 71, 78, 79] Subject received several different NACT regiments. They report higher baseline total hemoglobin concentrations ([HbT]) values and the greatest decrease in [HbT] levels for pCR and near-pCR subjects.[58]

Pakalniskis et. al. also looked at optical changes during NACT in 11 patients who received 4 different NACT regimens using a frequency domain DOT system.[57] They demonstrated a longitudinal decrease in total hemoglobin levels in pCR subjects but not non-pCR subjects. More recently Jiang et. al., looked a 19 patients that received 5 different therapies, including a subject on Avastin.[80] Employing a full-tomographic frequency domain system they reconstructed images of [HbT], SO_2%, and water fraction (WF) of the tumor bearing breast. Subjects that had a pCR showed an early percent

decrease in [HbT]. Several groups[26, 57, 59, 72-81] observed that [HbT] levels significantly decreases with pCR (p=0.001)[76,74]; and changes can occur as early as 1 week after starting NACT, with tumor Hb levels dropping 33% +/-7% in responders vs. non-responders (18% +/- 10%: z=0.008).[76]

Falou et al[86] observed 15 patients using diffuse optical spectroscopy for five time points during treatment. There were various chemotherapy regimens employed including chemoradiation for two subjects. One week after treatment they observed an initial increase in deoxy-hemoglobin concentrations ([Hb]), oxy-hemoglobin concentrations ([HbO_2]), WF, [HbT], and tissue optical index in responding patients, followed by a drop the whole breast. Conversely, poor response patients showed an initial decrease one week after treatment initiation followed by an increase optical parameters. A somewhat different observation was made by Soliman et. al. and Falou et. al. Employing DOS, they found an initial increase followed by a decrease in [HbT] during the first four weeks of treatment in patients that achieved a pCR after 6 months of NACT.[77, 86]

However, all of these previous studies have included patients on multiple different heterogeneous NACT regimens. [57-60, 73, 74, 77, 79-81, 86] Some studies even included chemoradiation therapy. [77, 86] Using different treatment regimens for patients in the same cohort may confound some of the obervations. For example, different drugs have differing anti-angiogenic effects and may affect optical signals in disparate ways. Only recently, Schaafsma et. al.[87] reported on results found in 22 patients who were all treated with the same NACT protocol. All patients received six cycles of docetaxel, doxorubicin, and cyclophosphamide (TAC) with or without zoledronic acid. Using a non-tomographic diffuse optical spectroscopy system they investigated the tumor bearing breast. They

found a statistically significant difference in the decrease in oxy-hemoglobin and deoxy-hemoglobin concentrations after the first cycle of NACT between those that achieved a pCR versus those with a partial response. No response subjects had an overall percent increase in hemoglobin concentration even to the mid treatment time point. Only 4 patients had a pCR. While results are promising, the low enrollment of complete responders makes it less significant.[87]

The study presented in this chapter differs from all previous studies in 3 major aspects. First, all patients follow the same NACT regimen, a standard taxane/anthracycline-based therapy. Second, the number of patients enrolled for this project (40) is one of the largest number of patients to be evaluated in an optical imaging NACT study to date. Third, our optical imaging system performs tomographic measurements on both breasts simultaneously, allowing for a direct comparison of the tumor bearing breast and the non-tumor bearing breast of the same patients.

Since there is sufficient evidence to suggest early changes in optical properties during NACT, we designed a pilot study to monitor tumor progression at six time points throughout treatment. We observed 40 women who all underwent the same chemotherapy regimen using a dual breast DOT system. We examined three dimensional tomographic images of both the tumor-bearing and contralateral breast. We quantified and analyzed the progression of the tumor over time and correlated chromophore concentrations to the RCB score of each patient.

4.2 Methods

We employed methods found in chapter 3 for the data pre-processing and analysis in this chapter. Briefly, we looked at five parameters: [HbO_2], [Hb], WF, total hemoglobin ([HbT] = [HbO_2] + [Hb]), and oxygen saturation (SO_2% = [HbO_2]/ ([HbO_2] + [Hb])). The normalized values were divided by the parameter values for the healthy breast ($[HbO_2]_{T/H}$, $[Hb]_{T/H}$, etc.) and divided by the weighted average for the non-tumor region in the tumor bearing breast ($[HbO_2]_{T/nT}$, $[Hb]_{T/nT}$, etc.). Additionally, the percent change from the baseline time point for each of the above mention parameters.

4.3 Results

4.3.1 Patient Summary

There were forty women recruited and imaged for the study with four subjects being excluded due to drop outs or machine malfunctions. Table 2.2 show a subject summary of the patients and their pathological response. Also, Table 2.3 show a summary of the number of subjects that were included in the analysis for each week for each response group.

4.3.2 Case Studies

The first subject is a 66 year old post-menopausal woman with poorly differentiated invasive ductal carcinoma. Her baseline tumor had a maximum clinical exam dimension of 2.3 cm. She had an RCB of 0. Her DOT images show a clear enhancement at the baseline time point in the right tumor bearing breast (Figure 4.1A). Yet, after the baseline imaging, there are no clear enhancements in either breast. The normalized quantified data shows a steep decrease by week 2 for $[HbO_2]_{(T/H)}$, $[HB]_{(T/H)}$, and $[HbT]_{(T/H)}$ (Figure 4.1C). SO_2% shows a slight decrease by week 2, but WF does not show much change

throughout the treatment. After the beginning of AC there seems to be another decrease in $[HBO2]_{(T/H)}$, $[HB]_{(T/H)}$, and $[HbT]_{(T/H)}$. The clinical exam measurement correlate well with the change in most of the parameters in which the tumor is no longer palpable by week 4.

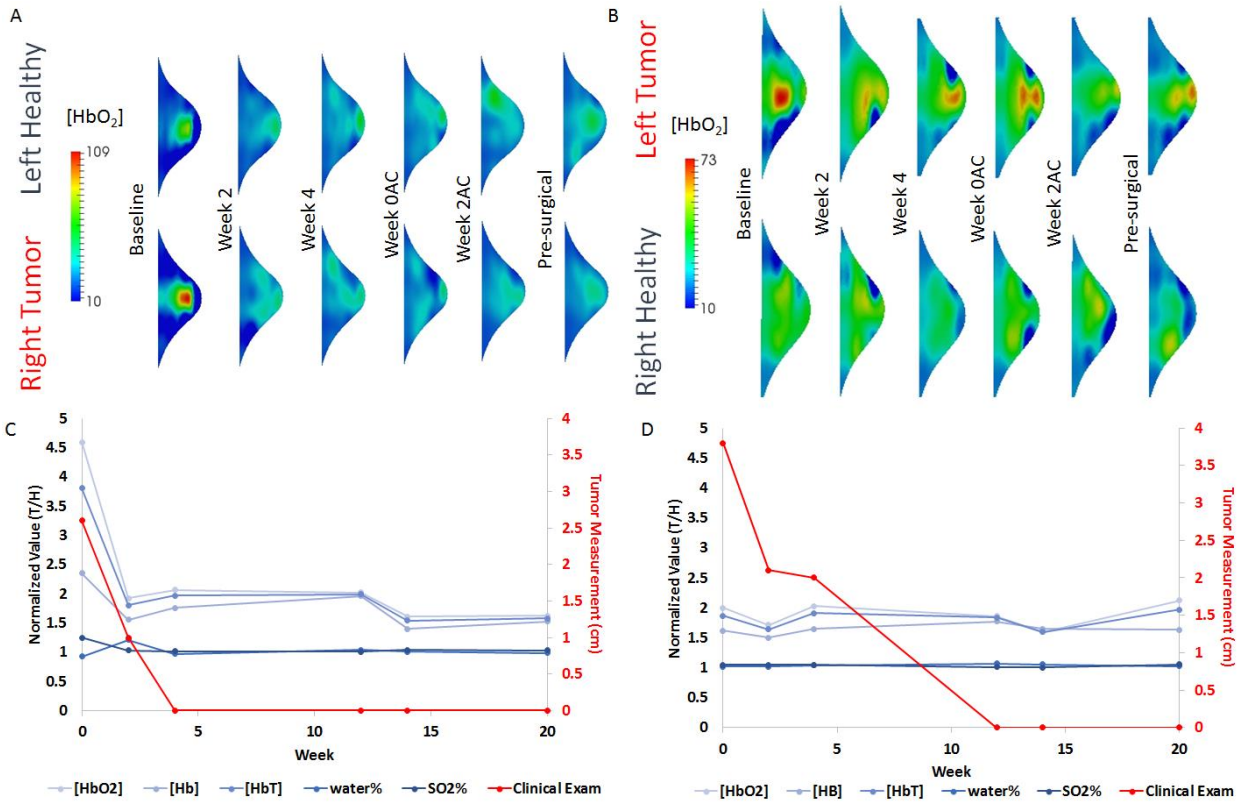


Figure 4.1 Two case studies of a subject with an RCB of 0 and a subject with an RCB of 2.7 (partial response). The DOT images for sagittal slices of the tumor bearing and contralateral breast in for the pCR subject (A) and the subject with the partial response (B). Quantified data showing the normalized value for each time point with the clinical exam tumor measurements for both the pCR (C) and partial responding subject (D).

The second subject is a 46 year old pre-menopausal subject with poorly differentiated invasive ductal carcinoma. Her baseline tumor was 3.8 cm based on clinical exam measurements. She had a partial response to treatment with an RCB 2.7. The DOT images for this subject show an enhancement in the left tumor bearing breast that is visible throughout treatment (Figure 4.1B). The quantified data reveals an increase in

all the parameters at the second time point (Figure 4.1D). Beyond week 2 the value of each of the parameters remain steady. The clinical exam measurements show that the tumor is no longer palpable by W0AC, but the subject still showed traces in the tumor during the AC portion of the treatment. The remaining excised tumor was still 1.4 cm in diameter.

4.3.3 Optical features during NACT

Table 4.1 shows a summary of of each feature for pCR and non-pCR subjects for each imaging time point. The p-value for each t-test performed between the subject groups is shown and highlighted in yellow if there is statistical significance. Here, we identify what features will that assist in determining the pCR subjects. The baseline and week 2 time points show the most features with statistical significance between response groups.

For baseline imaging, the more promising features that showed statistical significance between subject groups were $[HbO_2]_{(T/H)}$, $SO_2\%_{(T/H)}$, and $SO_2\%_{(T/nT)}$. The pCR group showed higher normalized $[HbO_2]_{(T/H)}$ and $SO_2\%_{(T/H)}$ compared to non-pCR with statistical significance of $p = 0.03$ and $p = 0.02$, respectively. For normalized $SO_2\%_{(T/nT)}$ pCR subjects were 1.12 ± 0.1 and non-pCR subjects were 1.06 ± 0.05 with a p-value of 0.03 ($p= 0.04$ for WT) between the groups (Figure 4.2).

Table 4.1 Summary of features for each time point and the p-value for the t-test between pCR and non-pCR subjects.

		Tumor Region			Normalized (T/H)			Normalized (T/nT)		
Time point	Parameter	pCR	non-pCR	p	pCR	non-pCR	p	pCR	non-pCR	p
Baseline	$[HbO_2]$	56.07± 10.64	48.84± 14.94	0.13	2.84 ±0.80	2.20± 0.80	0.03	2.60± 0.61	2.18± 0.72	0.08
	[Hb]	19.93± 4.05	19.30± 6.08	0.73	2.00 ±0.38	1.97± 0.89	0.89	1.87± 0.35	1.81± 0.59	0.74
	WF	0.53± 0.07	0.54± 0.04	0.82	1.04 ±0.10	1.02± 0.09	0.45	1.04± 0.02	1.03± 0.02	0.19
	[HbT]	75.90± 13.61	67.82± 17.51	0.15	2.55 ±0.63	2.12± 0.72	0.07	2.35± 0.49	2.05± 0.61	0.14
	SO_2%	72.22± 6.47	70.41± 4.90	0.35	1.11 ±0.11	1.04± 0.07	0.02	1.12± 0.10	1.06± 0.05	0.03
Week 2	$[HbO_2]$	34.68± 12.43	51.65± 16.32	0.003	1.81 ±0.93	2.40± 0.88	0.07	1.68± 0.53	2.31± 0.87	0.02
	[Hb]	16.58± 6.21	20.05± 6.53	0.14	1.57 ±0.56	2.01± 0.69	0.06	1.55± 0.38	1.95± 0.63	0.05
	WF	0.52± 0.05	0.53± 0.04	0.52	1.02 ±0.12	1.01± 0.08	0.88	1.02± 0.02	1.04± 0.03	0.02
	[HbT]	49.39± 16.11	70.91± 21.06	0.004	1.66 ±0.72	2.25± 0.80	0.04	1.57± 0.44	2.18± 0.78	0.02
	SO_2%	68.40± 3.60	71.19± 3.87	0.04	1.08 ±0.06	1.07± 0.06	0.45	1.07± 0.04	1.06± 0.05	0.90
Week 4	$[HbO_2]$	41.98± 14.98	41.97± 13.34	1.00	2.10 ±0.40	2.03± 0.71	0.75	2.07± 0.39	2.05± 0.56	0.90
	[Hb]	19.36± 7.08	16.87± 4.50	0.22	1.82 ±0.65	1.79± 0.55	0.90	1.81± 0.48	1.69± 0.37	0.39
	WF	0.52± 0.03	0.56± 0.05	0.07	0.98 ±0.09	1.02± 0.08	0.24	1.02± 0.02	1.02± 0.01	0.90
	[HbT]	60.31± 20.01	58.62± 16.73	0.79	1.96 ±0.44	1.95± 0.62	0.94	1.94± 0.33	1.92± 0.46	0.85
	SO_2%	67.60± 3.26	69.24± 4.56	0.28	1.09 ±0.08	1.03± 0.06	0.01	1.07± 0.05	1.06± 0.05	0.63
Week 0AC	$[HbO_2]$	33.82± 8.61	42.66± 16.15	0.12	1.80 ±0.32	2.15± 1.00	0.29	1.84± 0.34	2.08± 0.83	0.39
	[Hb]	17.36± 6.46	18.08± 8.03	0.81	1.73 ±0.49	1.95± 1.10	0.55	1.67± 0.45	1.84± 0.79	0.52
	WF	0.53± 0.05	0.53± 0.03	0.71	1.02 ±0.12	1.00± 0.08	0.59	1.01± 0.01	1.02± 0.02	0.13
	[HbT]	49.73± 12.05	60.47± 22.36	0.17	1.73 ±0.32	2.07± 0.96	0.29	1.73± 0.33	1.99± 0.77	0.31
	SO_2%	66.32± 5.27	69.77± 4.99	0.09	1.03 ±0.04	1.04± 0.06	0.70	1.06± 0.03	1.05± 0.04	0.39
Week 2AC	$[HbO_2]$	38.18± 9.33	41.24± 14.55	0.53	2.04 ±0.65	1.92± 0.68	0.65	2.02± 0.65	1.90± 0.59	0.61
	[Hb]	20.65± 7.54	16.44± 5.85	0.09	2.10 ±1.06	1.62± 0.49	0.10	1.95± 0.67	1.66± 0.48	0.16
	WF	0.52± 0.03	0.55± 0.05	0.15	1.01 ±0.04	1.03± 0.12	0.60	1.02± 0.01	1.02± 0.02	0.85
	[HbT]	56.98± 12.75	57.03± 18.09	0.99	1.99 ±0.66	1.81± 0.58	0.43	1.94± 0.60	1.81± 0.52	0.53
	SO_2%	67.08± 3.05	70.55± 3.37	0.01	1.05 ±0.07	1.06± 0.06	0.71	1.06± 0.05	1.05± 0.04	0.63
Pre-surgical	$[HbO_2]$	39.73± 7.55	40.39± 9.97	0.88	2.15 ±0.60	1.95± 0.56	0.43	1.80± 0.35	1.85± 0.50	0.82
	[Hb]	14.72± 1.94	16.08± 4.49	0.45	1.64 ±0.28	1.66± 0.42	0.93	1.52± 0.14	1.60± 0.40	0.61
	WF	0.54± 0.04	0.54± 0.04	0.99	1.01 ±0.08	1.02± 0.09	0.82	1.02± 0.01	1.02± 0.02	0.90
	[HbT]	54.15± 7.23	56.44± 13.24	0.67	1.98 ±0.49	1.85± 0.47	0.54	1.71± 0.28	1.77± 0.45	0.74
	SO_2%	71.67± 4.32	70.37± 3.32	0.42	1.08 ±0.07	1.04± 0.06	0.15	1.05± 0.02	1.04± 0.04	0.45

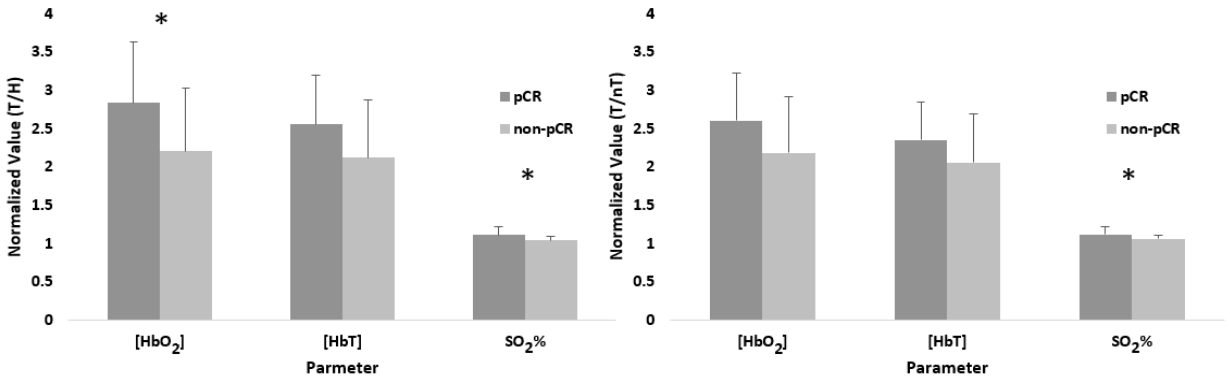


Figure 4.2 Baseline imaging comparing values for $[HBO_2]_{(T/H)}$, $[HBT]_{(T/H)}$, and $SO_2\%_{(T/H)}$ between pCR and non-pCR subjects (Left). Similarly baseline values for $[HBO_2]_{(T/nT)}$, $[HBT]_{(T/nT)}$, and $SO_2\%_{(T/nT)}$ (Right). Asterisk denotes statistical significance ($p<0.05$).

As for week 2 imaging, $[HbO_2]$, [HbT], $SO_2\%$, $[HbT]_{(T/H)}$, $[HbO_2]_{(T/nT)}$, $[Hb]_{(T/nT)}$, $WF_{(T/nT)}$, and $[HbT]_{(T/nT)}$ were all statistically significant between response groups. For each of these aforementioned DOT parameters, the pCR group had lower concentrations compared to the non-pCR group (Table 4.1). Figure 4.3 shows the three features at the week 2 time point with the lowest p-values: $[HbO_2]$, [HbT], and normalized $WF_{(T/nT)}$. Each of these parameter had higher values for the non-pCR group compared to the pCR subjects.

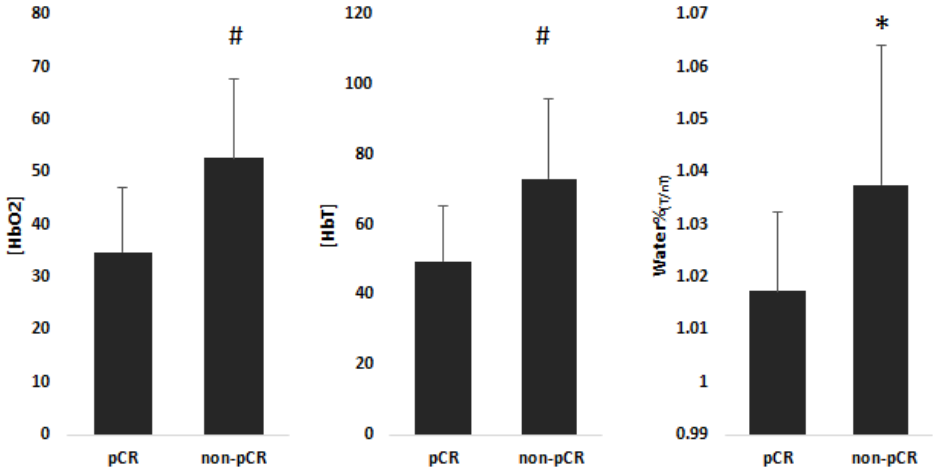


Figure 4.3 Comparison of pCR and non-pCR subjects at week 2 using imaging parameters that showed the highest correlation to RCB score in Table 3. An asterisk (*) denotes $p< 0.05$ and (#) denotes $p < 0.01$.

There were two other time points besides baseline and week 2 that showed statistical significance in a parameter. At week 4 $SO_2\%_{(T/H)}$ showed statistical significance between pCR and non-pCR subjects (p=0.01). Also, at week 2AC $SO_2\%$ showed statistical significance with non-pCR subjects having higher $SO_2\%$ levels than pCR subjects.

4.3.4 Baseline and Week 2 Correlation Coefficient

Table 4.2 Summary of correlation coefficients, and p-values of the linear regression of the RCB scores versus the baseline and week 2 parameters in the tumor region and normalized (T/H and T/nT). Highlighted in yellow are the statistically significant p-values.

Time Point	Parameter	Tumor Region		Normalized(T/H)		Normalized (T/nT)	
		R	p-value	R	p-value	R	p-value
Baseline	$[HbO_2]$	-0.29	0.09	-0.36	0.03	-0.33	0.052
	[Hb]	-0.11	0.53	-0.03	0.85	-0.08	0.62
	Water%	-0.01	0.97	-0.07	0.67	-0.25	0.14
	[HbT]	-0.28	0.10	-0.29	0.08	-0.28	0.10
	$SO_2\%$	-0.10	0.57	-0.31	0.07	-0.33	0.0504
Week 2	$[HbO_2]$	0.50	0.003	0.32	0.06	0.41	0.02
	[Hb]	0.15	0.40	0.21	0.23	0.23	0.18
	Water%	0.18	0.31	0.04	0.80	0.46	0.007
	[HbT]	0.46	0.006	0.34	0.05	0.41	0.02
	$SO_2\%$	0.36	0.04	0.02	0.92	0.07	0.68

Since most parameters showed the greatest statistical significance at baseline and week 2 we looked at the correlation coefficient for these time points. The parameters for the tumor region and normalized tumor region were correlated to the continuous RCB score and the statistical significance of the correlation was calculated. For baseline imaging the correlation that was the highest was the $[HbO_2]_{(T/H)}$. This had a negative

correlation coefficient which suggests that subjects that will have a low RCB or a pCR have higher $[HbO_2]_{(T/H)}$.

All features in for the baseline imaging had negative slopes and the all the features for week 2 had positive slopes. The feature $[HbO_2]_{(T/H)}$ had the only statistically significant correlation at baseline. In Figure 4.4 the linear regression analysis for RCB score versus $[HbO_2]_{(T/H)}$ is shown for both baseline and week 2 imaging. For baseline there is a negative correlation and for week 2 there is a positive correlation between RCB score and $[HbO_2]_{(T/H)}$.

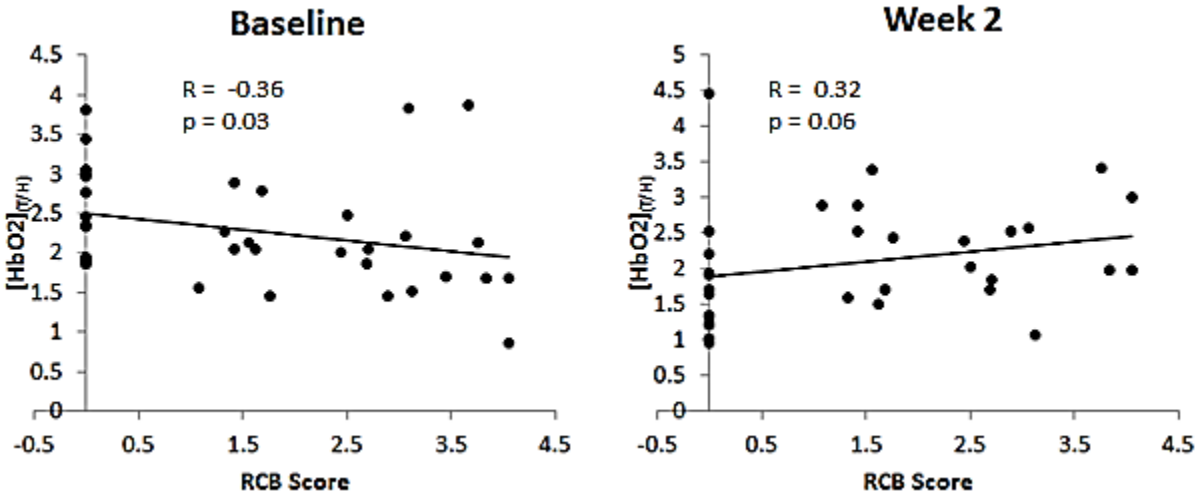


Figure 4.4 Linear regressions of the RCB score versus $[HbO_2]_{(T/H)}$ at baseline imaging (left) and week 2 (right).

4.3.5 Percent change from Baseline Imaging

After observing the absolute and normalized parameters, we looked at how these parameters change over time with respect to the baseline. In Table 4.3, the percent change (%Δ) of each parameter from the baseline imaging is shown. Here we can observe not only statistical significance between pCR and non-pCR subjects for some

features at the week 2 point but for all other time points as well, with the exception of the pre-surgical time point. However, week 2 and week 0AC have the most parameters with statistical significance.

Table 4.3 Summary of the percent change of each parameter from the baseline and the statistical analysis of pCR versus non-pCR subjects.

		Tumor Region			Normalized (T/H)			Normalized (T/nT)		
Time point	Parameter	pCR	non-pCR	p	pCR	non-pCR	p	pCR	non-pCR	p
Week 2	%Δ[HbO_2]	-36.44 ± 23.54	19.63 ± 69.58	0.01	-35.22 ± 30.91	22.94 ± 62.96	0.002	-33.48 ± 25.46	16.23 ± 52.63	0.003
	%Δ[Hb]	-14.31 ± 30.64	14.76 ± 49.24	0.07	-20.68 ± 25.16	16.45 ± 49.12	0.01	-13.64 ± 29.00	18.14 ± 45.64	0.03
	%ΔWF	-0.18 ± 14.37	-0.06 ± 9.13	0.98	-1.34 ± 15.29	-0.10 ± 11.86	0.63	-2.41 ± 1.85	0.59 ± 3.29	0.01
	%Δ[HbT]	-33.30 ± 22.36	16.12 ± 62.42	0.01	-34.03 ± 27.44	18.93 ± 58.11	0.003	-31.43 ± 23.32	15.33 ± 50.14	0.004
	%ΔSO_2%	-4.64 ± 9.19	0.99 ± 6.68	0.05	-2.17 ± 10.42	2.67 ± 7.19	0.08	-4.56 ± 7.16	-0.17 ± 4.20	0.03
Week 4	%Δ[HbO_2]	-25.63 ± 25.75	0.91 ± 73.97	0.24	-25.60 ± 19.66	6.59 ± 66.61	0.11	-19.74 ± 25.19	5.98 ± 54.71	0.13
	%Δ[Hb]	-2.58 ± 31.22	-4.36 ± 48.70	0.91	-9.26 ± 33.54	1.98 ± 50.88	0.58	-3.16 ± 31.01	-0.61 ± 36.25	0.84
	%ΔWF	-3.20 ± 5.53	4.15 ± 11.74	0.05	-3.25 ± 13.22	0.84 ± 11.69	0.32	-1.72 ± 2.02	-1.09 ± 2.72	0.48
	%Δ[HbT]	-20.97 ± 24.90	-1.64 ± 66.33	0.34	-23.32 ± 20.64	4.50 ± 62.89	0.16	-17.38 ± 22.35	3.13 ± 49.57	0.19
	%ΔSO_2%	-6.60 ± 8.17	-1.39 ± 7.10	0.06	-1.96 ± 12.09	-0.78 ± 7.21	0.56	-3.83 ± 10.59	-0.18 ± 4.35	0.16
Week 0AC	%Δ[HbO_2]	-37.81 ± 19.76	-6.79 ± 38.72	0.02	-33.24 ± 21.70	0.04 ± 36.19	0.01	-27.66 ± 19.06	-2.27 ± 30.40	0.02
	%Δ[Hb]	-9.19 ± 36.46	-6.14 ± 27.71	0.80	-12.87 ± 25.66	-0.55 ± 27.02	0.38	-12.12 ± 30.73	0.90 ± 22.25	0.19
	%ΔWF	-1.89 ± 7.91	-1.13 ± 9.42	0.83	-2.05 ± 11.49	-2.13 ± 10.88	0.97	-2.55 ± 1.74	-0.73 ± 1.95	0.02
	%Δ[HbT]	-32.32 ± 20.42	-8.15 ± 32.52	0.04	-29.84 ± 20.21	-1.51 ± 30.87	0.02	-26.71 ± 16.79	-2.06 ± 26.46	0.01
	%ΔSO_2%	-8.86 ± 11.89	-0.30 ± 10.06	0.05	-7.92 ± 10.25	0.45 ± 8.41	0.02	-3.98 ± 8.62	-1.43 ± 3.85	0.26
Week 2AC	%Δ[HbO_2]	-26.70 ± 31.86	-5.38 ± 48.26	0.20	-20.58 ± 37.22	-2.23 ± 39.60	0.17	-19.30 ± 30.50	-4.99 ± 36.24	0.28
	%Δ[Hb]	8.55 ± 49.15	-5.71 ± 39.99	0.39	10.17 ± 66.68	-4.31 ± 36.28	0.39	3.30 ± 38.43	0.72 ± 33.66	0.85
	%ΔWF	-4.71 ± 6.22	2.98 ± 12.39	0.06	-4.02 ± 8.99	0.83 ± 15.62	0.34	-1.86 ± 1.91	-1.11 ± 2.38	0.38
	%Δ[HbT]	-19.46 ± 32.74	-7.76 ± 42.00	0.43	-14.48 ± 41.87	-4.99 ± 35.07	0.47	-15.74 ± 29.92	-4.67 ± 32.45	0.36
	%ΔSO_2%	-7.08 ± 9.03	0.40 ± 8.73	0.03	-5.95 ± 7.43	1.89 ± 7.99	0.008	-4.24 ± 6.52	-1.25 ± 3.44	0.10
Pre-Surgical	%Δ[HbO_2]	-20.58 ± 29.34	0.69 ± 76.27	0.48	-13.71 ± 39.87	7.88 ± 68.13	0.42	-28.44 ± 23.88	-0.25 ± 60.54	0.25
	%Δ[Hb]	-19.03 ± 18.27	-2.34 ± 62.01	0.49	-10.91 ± 25.15	2.28 ± 56.59	0.65	-22.05 ± 15.64	1.30 ± 50.73	0.25
	%ΔWF	-0.04 ± 7.95	0.83 ± 9.28	0.83	-1.88 ± 12.55	-1.14 ± 11.09	0.72	-1.73 ± 2.67	-1.21 ± 2.71	0.66
	%Δ[HbT]	-20.50 ± 24.52	-1.74 ± 70.05	0.50	-13.18 ± 34.90	4.69 ± 62.78	0.51	-26.71 ± 21.26	-0.06 ± 56.84	0.24
	%ΔSO_2%	0.49 ± 5.96	0.54 ± 10.09	0.99	-2.31 ± 9.58	0.76 ± 9.22	0.35	-3.26 ± 4.41	-2.41 ± 3.97	0.64

The non-normalized percent change tumor measurements had the most statistically or near statistically significant features. Figure 4.5 presents the features that were statistically significant at least at the week 2 time point. Each features shows an increase in the percent change for the non-pCR subjects and a percent decrease for the pCR subjects at the week 2 time point. The %Δ[HbO_2] (Figure 4.5A) shows statistical

significance at only the week 2 and week 0AC time points. The pCR subjects seem to not change much after the initial decrease at week 2 and then there is a slight increase occurring after week 12 (week 0AC). The %Δ[Hb] shows no statistical significance beyond the second week of treatment (Figure 4.4B). The %Δ[HbT] (Figure 4.4C) had a similar trend as the %Δ[HbO_2]. Last, %ΔSO_2% shows statistical significance between groups for all time points with the exception of the pre-surgical time point. The pCR subjects show a decrease in %ΔSO_2%, which begins to increase again after AC began to be administered (week 12). The non-pCR subjects show that %ΔSO_2% slightly decreases by week 4 but continues to increase throughout the rest of treatment.

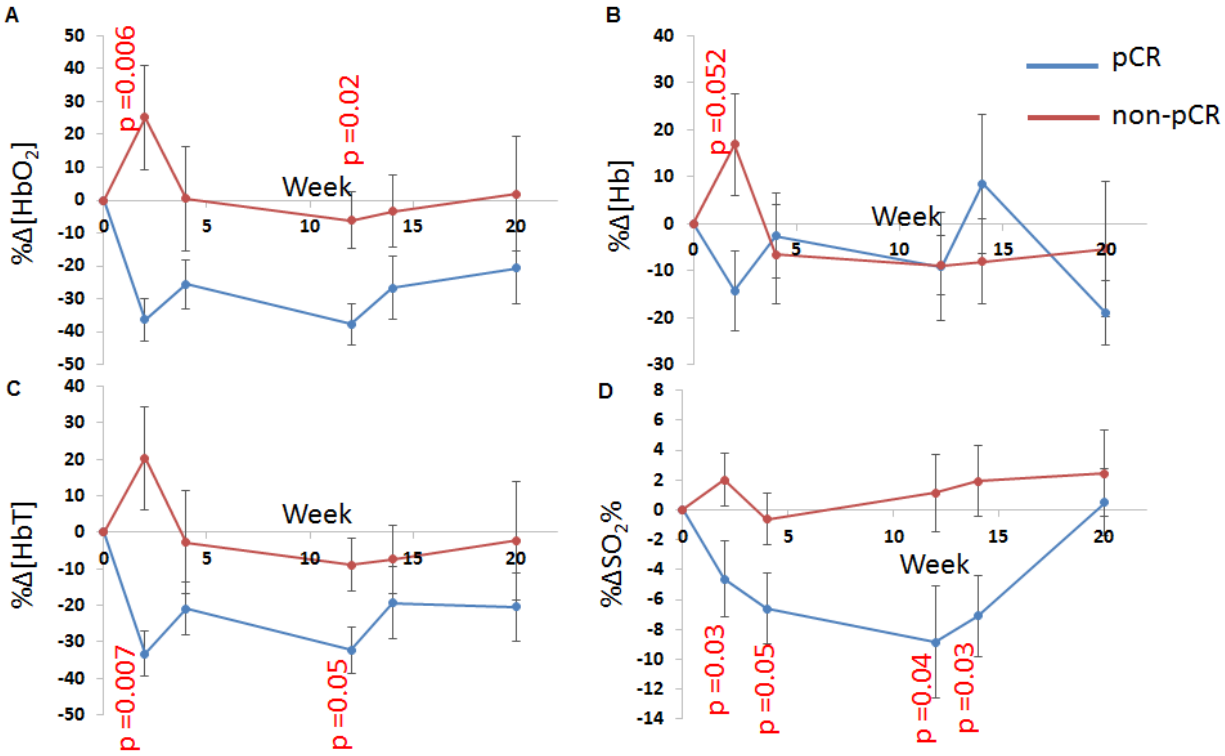


Figure 4.4 Longitudinal line plots of pCR and non-pCR subjects for %Δ[HbO_2] (A), %Δ[Hb] (B),%Δ[HbT] (C), and %ΔSO_2% (D).

At the week 2 time point, the percent change in $[HbO_2]_{(T/H)}$, $[Hb]_{(T/H)}$, and $[HbT]_{(T/H)}$ were statistically significant, as well as, all the features for the (T/nT) normalization scheme (Table 4.3). For each of these features the pCR group experienced a percent decrease, while non-pCR subjects demonstrated an increase in the normalized features. For the Week 0AC time point, the statistically significant features normalized to the healthy breast for the Week 0AC time points were the percent change in $[HbO_2]_{(T/H)}$, $[Hb]_{(T/H)}$, and $[HbT]_{(T/H)}$. As for the (T/nT) normalization scheme, in $[HbO_2]_{(T/nT)}$, $WF_{(T/nT)}$, and $[HbT]_{(T/nT)}$ showed statistical significance between the pCR and non-pCR groups. Last, the percent change in $SO_2\%_{(T/H)}$, showed statistical difference at the Week 2AC time point.

4.3.6 Correlation Coefficient of the Percent Change

Similar to the baseline and week 2 analysis, we observed the correlation coefficient for the percent change from the baseline imaging for all parameters. The percent change in the tumor region is significantly correlated with the RCB score for $\%\Delta[HbO_2]$, $\%\Delta[Hb]$, and $\%\Delta[HbT]$ for week 2. The $\%\Delta WF$ is correlated only for the week 4 and week 2AC time point (Table 4.4).

For the (T/H) normalized parameters, the strongest correlated parameters for week 2 are $\%\Delta[HbO_2]_{(T/H)}$, $\%\Delta[Hb_2]_{(T/H)}$, and $\%\Delta[HbT]_{(T/H)}$. However, beyond the second week, $\%\Delta[HbO_2]_{(T/H)}$ shows significant correlation at week 4 and the week 2AC time points. Additionally, $\%\Delta[HbT]_{(T/H)}$ has significant correlation at week 4, while $\%\Delta SO_2\%_{(T/H)}$ is strongly correlated at Week 2AC. Only $\%\Delta[Hb]_{(T/H)}$ in the pre-surgical time point showed statistical significance.

Table 4.4 Summary of linear regression analysis of RCB score versus the percent change from baseline imaging of each parameter for each time point.

Time Point	Parameter	Tumor Region		Normalized(T/H)		Normalized (T/nT)	
		R	p-value	R	p-value	R	p-value
Week 2	%Δ[HbO_2]	0.5	0.003	0.51	0.002	0.55	0.001
	%Δ[Hb]	0.37	0.03	0.43	0.01	0.34	0.051
	%ΔWF	0.1	0.59	0.04	0.83	0.52	0.002
	%Δ[HbT]	0.5	0.003	0.51	0.002	0.52	0.002
	%ΔSO_2%	0.27	0.12	0.28	0.11	0.4	0.02
Week 4	%Δ[HbO_2]	0.29	0.1	0.38	0.03	0.36	0.04
	%Δ[Hb]	0.11	0.53	0.24	0.16	0.14	0.45
	%ΔWF	0.34	0.05	0.13	0.46	0.21	0.24
	%Δ[HbT]	0.26	0.14	0.35	0.04	0.32	0.07
	%ΔSO_2%	0.19	0.28	0.07	0.69	0.3	0.09
Week 0AC	%Δ[HbO_2]	0.31	0.09	0.34	0.07	0.36	0.05
	%Δ[Hb]	0.11	0.57	0.25	0.18	0.25	0.18
	%ΔWF	0.13	0.48	0.03	0.85	0.34	0.06
	%Δ[HbT]	0.31	0.09	0.36	0.05	0.4	0.03
	%ΔSO_2%	0.22	0.22	0.29	0.12	0.19	0.29
Week 2AC	%Δ[HbO_2]	0.35	0.051	0.37	0.04	0.39	0.03
	%Δ[Hb]	-0.05	0.77	-0.06	0.76	0.06	0.74
	%ΔWF	0.51	0.004	0.2	0.27	0.3	0.10
	%Δ[HbT]	0.29	0.11	0.28	0.13	0.35	0.052
	%ΔSO_2%	0.3	0.10	0.44	0.01	0.32	0.08
Pre-surgical	%Δ[HbO_2]	0.29	0.16	0.35	0.08	0.4	0.04
	%Δ[Hb]	0.33	0.10	0.39	0.05	0.42	0.03
	%ΔWF	0.07	0.74	-0.03	0.88	0.17	0.41
	%Δ[HbT]	0.3	0.13	0.36	0.07	0.41	0.04
	%ΔSO_2%	-0.12	0.55	0.03	0.9	0.12	0.57

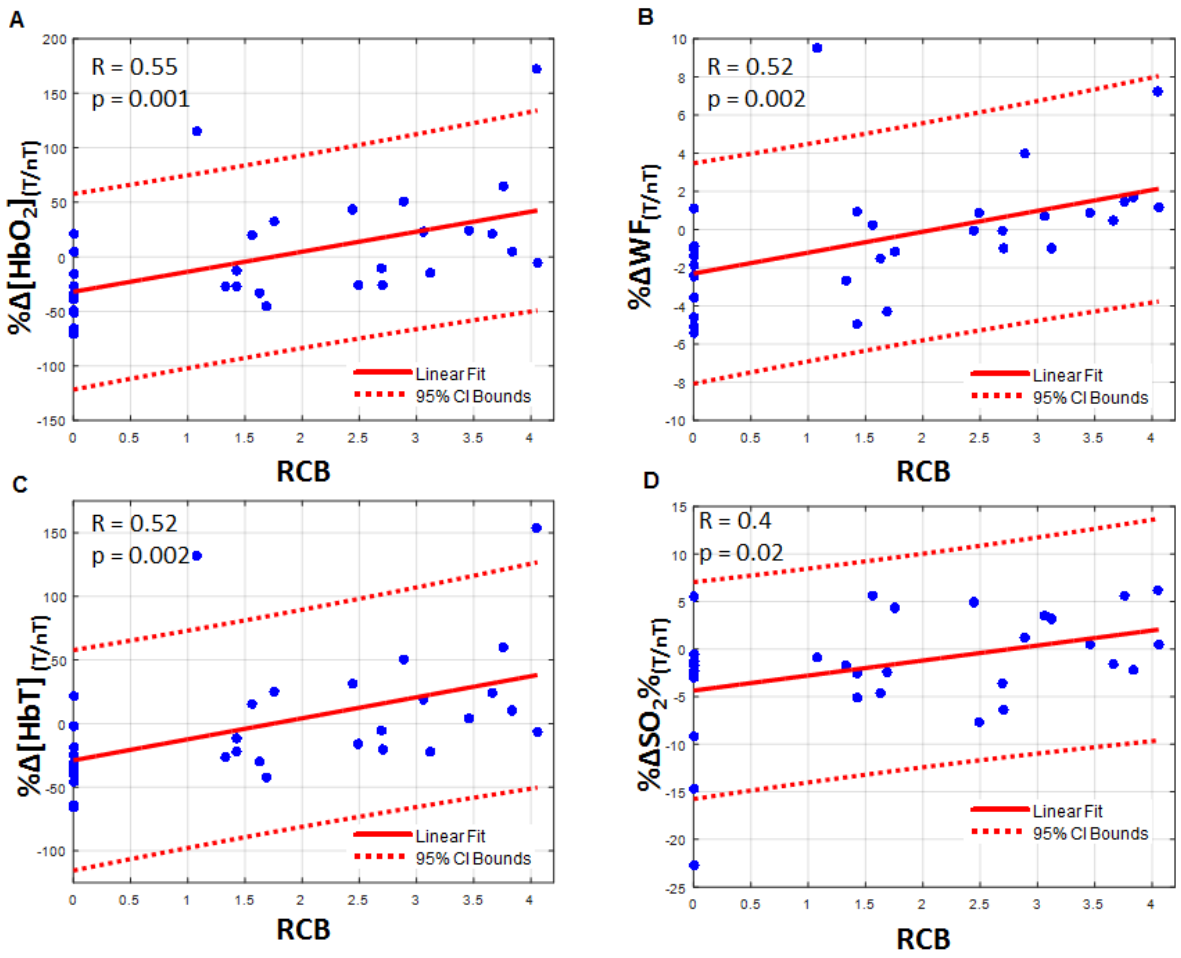


Figure 4.5 Linear regression of the RCB score versus the $\%\Delta[HbO_2]_{(T/nT)}$ (A), $\%\Delta WF_{(T/nT)}$ (B), $\%\Delta[HbT]_{(T/nT)}$ (C), $\%\Delta SO_2\%_{(T/nT)}$ (D). Solid lines show the linear fit to the data and the dotted lines show the 95% confidence intervals (CI) upper and lower bounds.

For the (T/nT) normalization scheme, all the parameters show high correlation at the week 2 time point. Figure 4.5 shows the linear correlation RCB score versus $\%\Delta[HbO_2]_{(T/nT)}$ (A), $\%\Delta WF_{(T/nT)}$, $\%\Delta[HbT]_{(T/nT)}$, and $\%\Delta SO_2\%_{(T/nT)}$, which all showed statistical significance at week 2. Each of these displayed features have a positive correlation compared to RCB score.

The features $\%\Delta[HBO2]_{(T/nT)}$ and $\%\Delta[HbT]_{(T/nT)}$ had statistical significance or near statistical significance for all the time points including the pre-surgical time point.

4.3.7 Longitudinal Trends

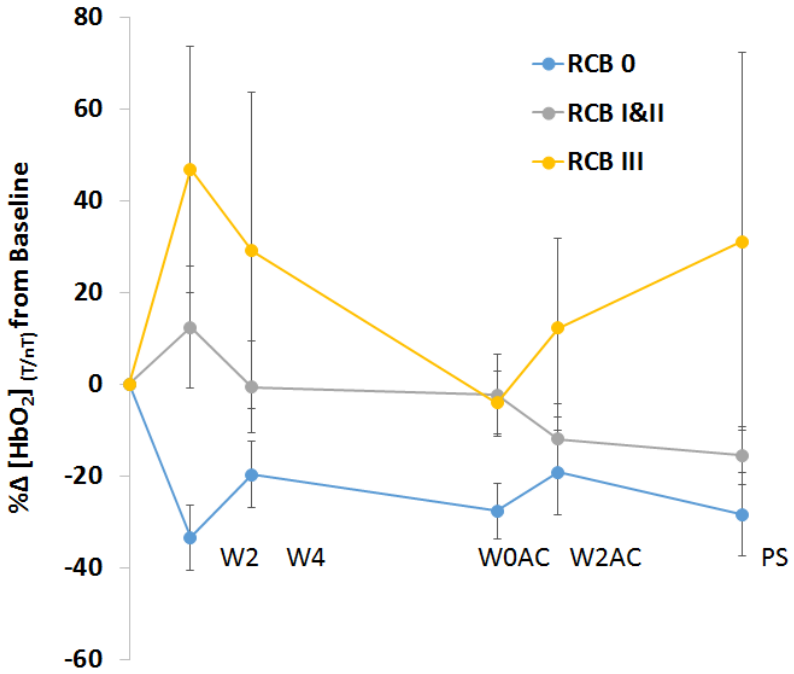


Figure 4.6 The percent change from baseline imaging of $[HBO_2]_{(T/nT)}$ over the course of NACT for RCB 0 (pCR) subjects, RCB I and II subjects, and RCB III subjects. Error bars represent standard error.

Since the $\%\Delta[HbO_2]_{(T/nT)}$ show the greatest correlation among the all the time points we looked at how these parameters change over the course of treatment. In figure 4.6, each subject group displays non-linear trends of how these parameters change over the course of treatment. The RCB 0 subjects show a decrease in $\%\Delta[HbO_2]_{(T/nT)}$ by week 2, but then the percent change from the baseline does not change much at the other time points. The RCB I&II group shows a slight percent increase by week 2, but does not seem to change much from the baseline. However, after the W0AC time point this group shows a percent decrease at the Week 2AC and pre-surgical time points. Last, the RCB III group increases about 47%, decreases back to baseline values by Week 0AC, and then shows about a 31% increase by the pre-surgical time point.

4.3.8 Predicting Tumor Response

ROC curves were created for each parameter and the five parameters with the highest Youden index for all imaging time points as seen in Table 4.5. These parameters are all during week 2 and consist of the percent change in either oxy-hemoglobin or total hemoglobin (normalized or otherwise). The parameter with the highest area under the curve (AUC) is $\%\Delta[HbO_2]_{(T/nT)}$, where we observed a sensitivity of 76.9% (10/13 subjects) and a specificity of 90.5% (19/21 subjects). Therefore, the overall accuracy using a cutoff value of -27.5% is 85.3% (29/34 subjects). The highest sensitivity was 100% (13/13 subjects) for $\%\Delta[HbO_2]$ for week 2. The specificity for this parameter was 61.9% (13/21), an accuracy of 76.5% (26/34 subjects), a PPV of 61.9% (13/21 subjects), and a NPV of 100% (13/13 subjects).

Table 4.5 Summary of parameters with the highest Youden index values for baseline and week 2 imaging.

Parameter	Time Point	Sensitivity	Specificity	Accuracy	AUC	PPV	NPV
$\%\Delta[HbO_2]_{(T/nT)}$	Week 2	76.9%	90.5%	85.3%	0.87	83.3%	86.0%
$\%\Delta[HbO_2]_{(T/H)}$	Week 2	76.9%	90.5%	85.3%	0.86	83.3%	86.0%
$\%\Delta[HbT]_{(T/nT)}$	Week 2	84.6%	76.2%	85.3%	0.82	68.8%	89.0%
$\%\Delta[HbT]_{(T/H)}$	Week 2	76.9%	85.7%	82.4	0.86	76.9%	86.0%
$\%\Delta[HbO_2]$	Week 2	100%	61.9%	76.5%	0.84	61.9%	100%

4.4 Discussion

4.4.1 Summary of Results

The week 2 time point shows the highest correlation for $[HbO_2]$ and [HbT] to the RCB score of the subject, even when the data is normalized or the percent change of the

normalized values are calculated. We found a statistical decrease in [HbT], which corresponds to the multiple studies done on optical NACT monitoring.[26, 48, 57-60, 72-81] Additionally, the baseline $SO_2\%_{(T/H)}$ values for pCR subjects were higher than non-pCR subjects. This is similar to the work done by Ueda et. al., where they found an increase in SO2% in subjects that would receive a pCR to NACT.[60]

However, there is statistically significant correlation between RCB score and the DOT parameters ***throughout*** the entirety of NACT. Therefore, while the week 2 and even baseline time points may be used to evaluate the efficacy of treatment, the other imaging time points can also be used to determine if the subject will respond to NACT as well. Determining response throughout therapy can be beneficial to monitor if the tumor is still progressing or not. Therefore, there can be multiple check points within the NACT to be able to monitor and assess treatment to assure that the patient has the optimal response to NACT.

4.4.2 Physiological Explanation

Taxanes cause cell death by inhibited cell growth, differentiation, and proliferation by stabilizing microtubules within the cells. Essentially taxanes cause an arrest in the G2 phase of cell division prevent further maturation and proliferation.[88, 89] There has also been evidence of taxanes inhibiting VEGF, which would suppress angiogenesis.[88, 90] Symmans et. al. demonstrated an increase in apoptosis and mitotic arrest in BC patient receiving paclitaxel within days of the first dose.[91] Additionally in terms of early vascular changes, there have been studies that looked at interstitial fluid pressure (IFP) within the breast tumors after taxane administration,[92] which is caused by the massive apoptosis of responding tumor cells.[93] The early changes that we have observed may be caused by

these physiological phenomenon. First, there was a decrease in [Hb] and [HbO2] by the second week of treatment for pCR subjects. This can be caused by the antiangiogenic effects that taxanes have on tumors. Also, the massive apoptosis of cancer cells could restore the microenvironment making the area less hypoxic and contributing to the decrease in [Hb]. Last, we noticed that there was a decrease for pCR subject in $WF_{(T/nT)}$ by week 2. This could be contributed to the decrease in IFP within the tumor region compared to the surrounding healthy tissue in the tumor bearing breast. Therefore, the early optical response of pCR subjects may possibly be contributed to the decrease in microvessels, hypoxia, and IFP.

Anthracyclines stop the process of replication by preventing DNA from resealing after transcription. Cyclophamide similarly works by disrupting the DNA to prevent cell growth and division.[89] Anthracycline has shown to increase apoptosis and decrease proliferation in breast cancer.[94] During this part of the treatment there seemed to be very little change to the pCR subject. However, the RCB I/II subjects began to decrease beyond baseline values in $[HBO_2]_{(T/nT),}$ while RCB III subjects demonstrated and increase in $[HBO_2]_{(T/nT)}$ after AC treatment (refer to fig 4.6). As a result, there seems to be no further change in pCR subject, the RCB I/II subjects seem to benefit from the change in chemotherapy agents, and the RCB III subjects seem to worsen.

4.4.3 Conclusion

Early imaging time points, as other studies have suggested, are still optimal for determining the response of the subject. Specifically we saw that the $\%\Delta[HbO_2]_{(T/nT)}$ at week 2 time point yielding the highest sensitivity and specificity for determining a pCR or not. Additionally, there is a positive statistically significant correlation between

$\%\Delta[HbO_2]_{(T/nT)}$ and RCB score for every time point after baseline imaging. This suggests that it may be possible to monitor tumor progression throughout NACT, not only initially predict response. Monitoring would be more beneficial in some patient cases. Hence, if a subject is predicted to respond by the second week, but further monitoring shows that she has not responded as well as originally perceived, a decision could be made on whether to proceed with prescribe chemotherapy agent or change drugs to help optimize the subject's outcome. These results are promising since it could lay the ground work for monitoring, as well as, tumor prediction.

Chapter 5: Dynamic Data Processing

5.1 Introduction

In a previous pilot study we have shown evidence that the hemodynamic response to a simple breath hold can be used to diagnose breast cancer. Tumors tend to have inherently different time traces compared to healthy tissue.[39] Here we hypothesize that vascular changes occur early in the treatment process and that related hemodynamic parameters can be used to predict NACT treatment outcome. The different features from the hemodynamic traces that we extracted from the dynamic DOT imaging will be correlated with the pathological response of the subject.

There are three steps that were necessary for dynamic DOT analysis: pre-processing, image quantification, and feature extraction.

5.2 Challenges to Overcome

There were several challenges to the pre-processing procedure that needed to be overcome. First, there needed to an efficient and quick method to automate the pre-processing procedure. Second, there were high frequency noise that made the breath hold trend difficult to determine. Third, each dynamic movie reconstruction took about 24 hours to complete, but not each breath hold was viable for analysis. Therefore, some breath holds needed to be reconstructed again, so there needed to be a procedure to determine before the reconstruction process to determine if the breath hold was usable. Additionally, some detector signals would cause issues with the reconstructions and would need to be removed, so there needed to be a process to determine these points.

Last, in order to fix these issues there needed to be a means of visualizing the raw data and the input information for the reconstructions to see how improvements could be made to the pre-processing procedure. What made data visualizing challenging was the enormous amount of data that we collect. If there are 32 sources, 64 detectors, 2 wavelengths, and a thousand frames then there are over 4 million data points to observe. Therefore, there were sufficient efforts made to develop techniques to visualize all of this data for pre-processing.

Image quantification also needed to be improved. Originally, for each frame of the reconstruction, the location of the maximum value was found and then a 1 cm diameter sphere around the maximum was averaged. Each average became a data point in time and the changes in the reconstruction could be viewed over the course of about a minute. One issue with this method is that the 1 cm sphere is free to move within the breast volume, so the maximum value could bounce from one region to the other, which would not occur in typical breast tumor physiology. Also, the tumor region usually is not confined to a 1 cm sphere volume or shape. Therefore, techniques were developed to quantify the data to determine a stationary volume that resembled the size and shape of the tumor.

For the feature extraction, we needed to determine physiological parameters that would assist in predicting tumor response. However, one problem was that some subjects could not hold their breath for the entire desired length of thirty second. Despite this challenge, if a subject held their breath for only fifteen seconds, half the time recommended, there were still clear hemodynamic effects to the breast tissue as seen from our optical imaging. Yet, this caused a large variation in the amount of time that the subjects could hold their breath. Therefore, certain parameters such as, time to peak or

a certain percentage, full width half maximum, and time to recovery could not be analyzed. As a result, we looked more a parameters that were not dependent on the breath hold length such as normalized peak values and washout rates.

5.3 Pre-processing

5.3.1 Overview

The pre-processing procedure before running a reconstruction is outlined in Figure 5.1. First, the raw data was pre-processed by applying a discrete cosine transform (DCT) filter that uses the first 15 coefficients to reduce high-frequency noise. Then one of the breath holds is chosen for reconstruction and the measurement files are created. Measurement files contain the filtered dynamic signal for each source-detector signal and the averaged baseline signal, which were calculated from the raw data. Next, the forward model is run in order to get the prediction measurement for the reconstruction. The input files are created using the measurement files and the prediction measurement. There may be an additional step in which poor detector signals are removed from the input files directly. Last, the reconstruction was run.

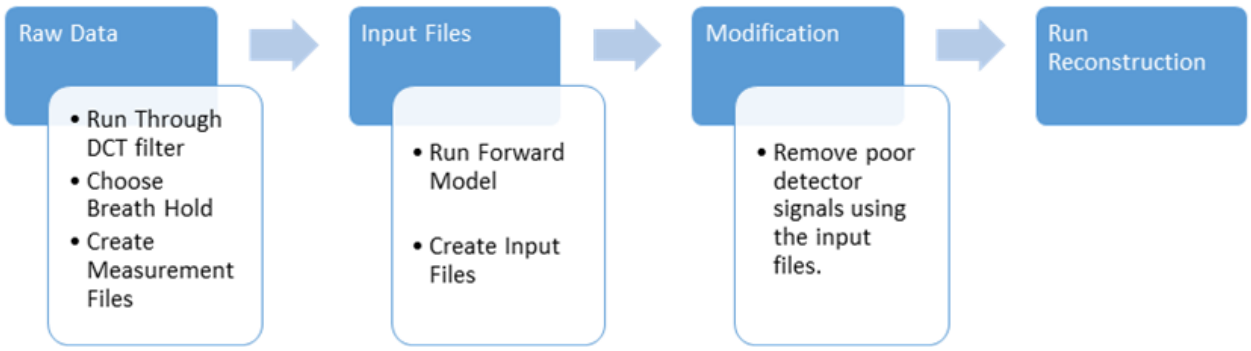


Figure 5.1 Flow diagram of how data is processed before the reconstructions are run.

Similar to the static input files, the input files are created using the following equation:

$$z(t)^{\lambda} = \frac{M(t)_{tar}^{\lambda}}{M_{ref}^{\lambda}} P^{\lambda}$$

where λ is the wavelength, z(t) is the input to the reconstruction for time t, M is a measurement, and P is the prediction measurement. $M(t)_{tar}$ represents the measurement at time t during the breath hold. M_{ref} is the baseline measurement taking as an average of the 30 seconds of data prior to the subject holding their breath. Then the signals that showed a signal to noise ratio of less than 10dB were labelled for removal from the reconstruction. These two steps gave the input files to the reconstruction.

5.3.2 Automated Program

In order to automate and quicken the pre-processing, a simple user interface was designed to guide the user easily through the necessary steps to pre-process the data. Figure 5.2 shows the first step where the measurement files are created. The program covers two stages of the pre-processing procedure. First, the raw data is filtered and the measurement files are created. Second, the calculated input files are used to remove any potentially erroneous signals.

The first step for creating the measurement files is loading the raw data. The raw data are kept in files that the program can automatically search through to obtain the desired patient and time point. Next the all the data is for each wavelength is displayed as a check to make sure there are no obvious issues with the data (i.e. missing wavelength data due to malfunction). Figure 5.3 shows an example of the signals from the raw data.

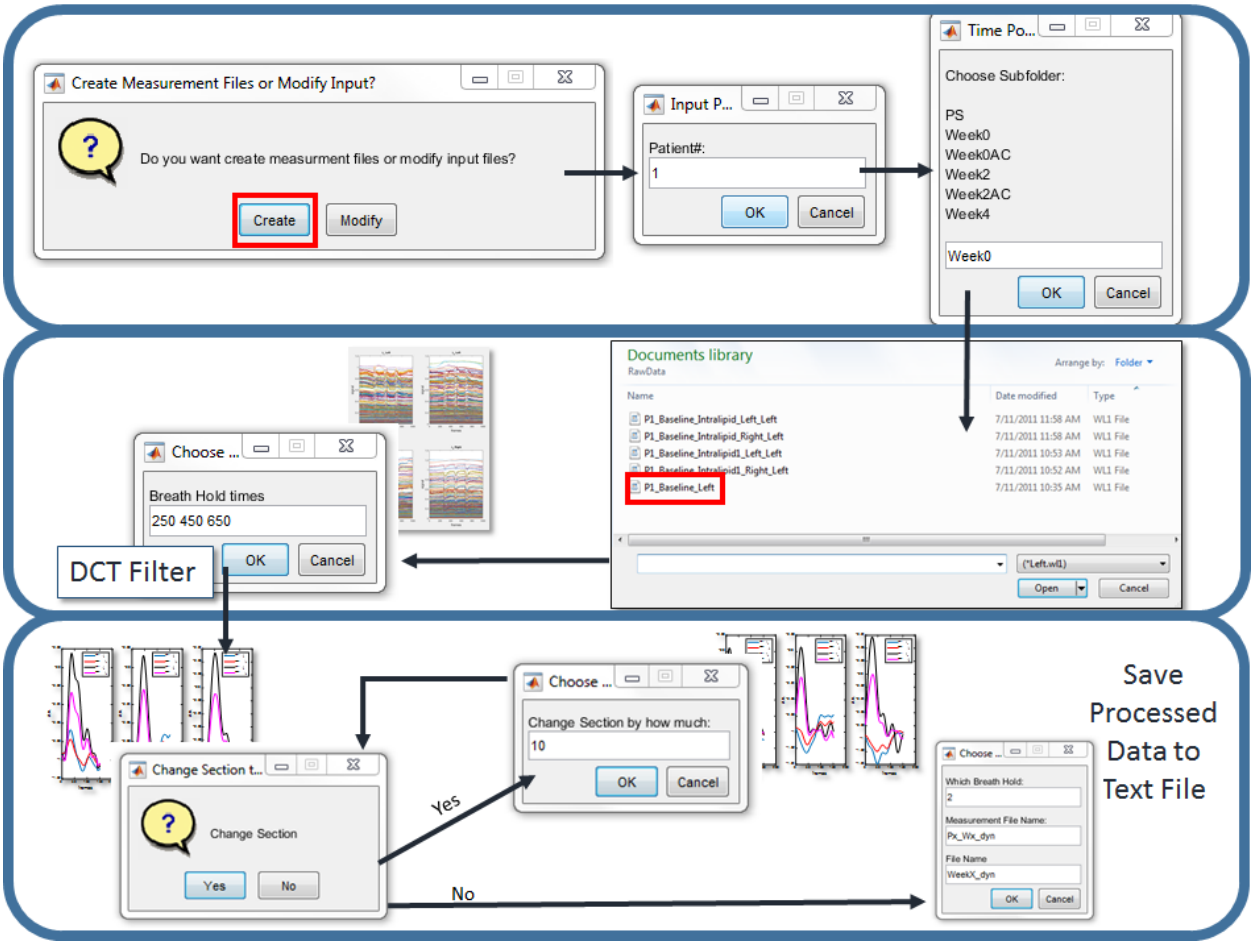


Figure 5.2 Flow diagram of MatLab program to create text files for reconstruction input.

Then the program asks for the starting frames for each breath hold. The program then performs DCT filter on each of these sections of data and displays the filtered signal (Figure 5.4). The filtered signal display shows the negative sum of all the detector signals. This has become an essential check in the dynamic reconstruction process. Before this check point was established, the optimal breath hold for reconstruction was difficult to establish. Since occasionally the patient might move after a breath hold or talk, there are sometimes motion artifacts or other issues with the breath hold. With this MatLab program, we can see each breath hold at once, and determine the best one for reconstruction. Otherwise, potentially, each breath hold might need to be reconstructed

to find the best one, which was time consuming since each reconstruction could take a day to complete.

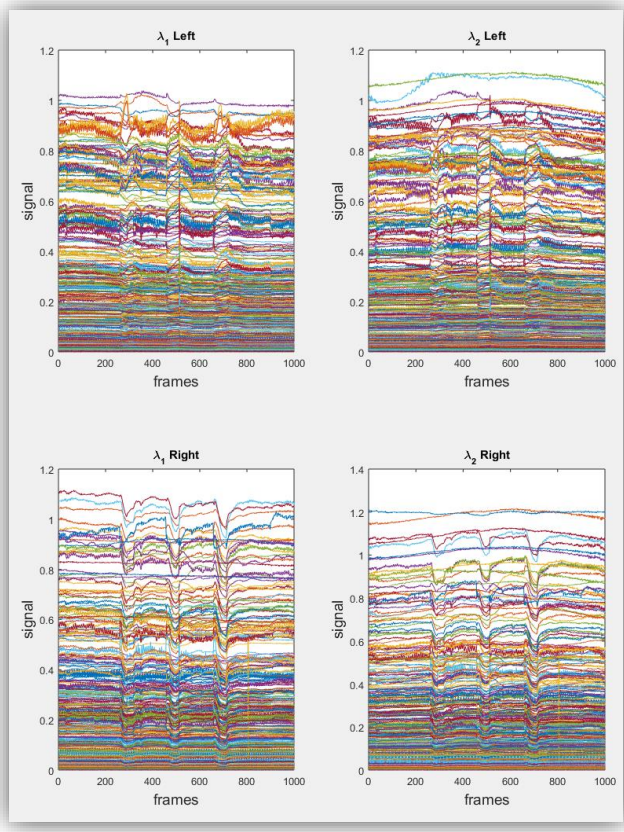


Figure 5.3 Visualization of each raw signal for an entire imaging from each source detector pair for each wavelength for the left and right breast.

There is an option to shift the breath hold frames one way or the other. Occasionally, the breath hold does not start right at the beginning of the frames that we instructed during the subject imaging. The subject could either start late (i.e. take their time beginning the breath hold) or start before we finished giving directions. Therefore, this option allows us to shift the data filtered a couple of frames in order to begin the reconstruction at the start of the breath hold. Once the breath hold is filtered and chosen,

the user can choose how to name the files and the text files are automatically saved to a folder.

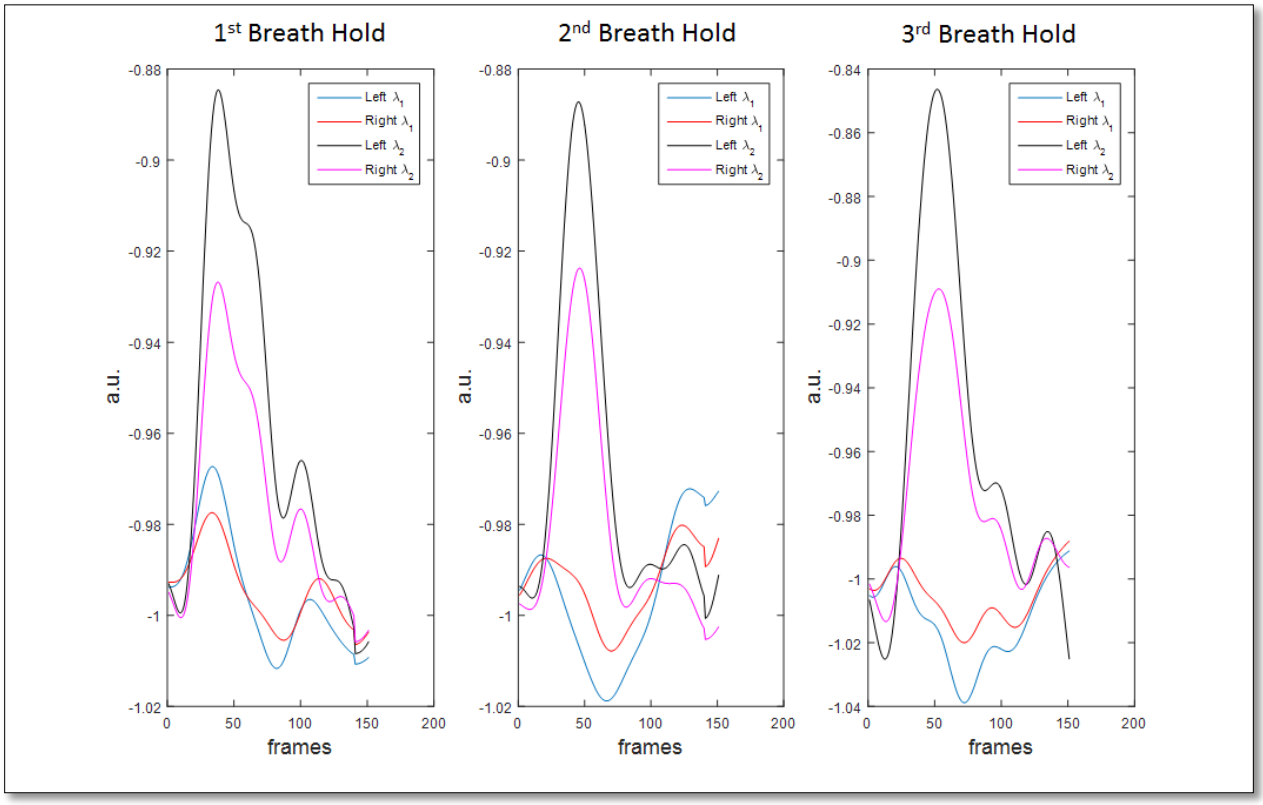


Figure 5.4 Time traces for each breath hold that shows the normalized summation of all the detectors for each wavelength for each breast.

5.3.3 Modify Input Files

The other option for this program is to modify the input files. After the measurement files have been created and the forward model run, then the created input files can also be used to check for any errors in the data. Similarly to the "creating the measurement" process of the program, the program has a "modify input file" process, which also guides the user to correct files they wish to use. Once the files are loaded, each frame of the breath hold can be displayed. In Figure 5.5 is an example of one frame of data for a patient that has poor signal in one of the detectors. After the user is finished

looking through the frames, the program prompts whether to remove any sources or detectors from the reconstruction. From the example giving in Figure 5.5, detector 54 would need to be removed from the reconstruction, since it is showing a much lower signal than the rest of the data. This process is important since aberrant detectors can severely impair the quality of the dynamic reconstructions. As seen in Figure 5.6, with the poor detector used in the reconstruction input data, there is no visible breath hold and the magnitude is far above what is typically expected. With the detector removed, the breath hold is visible.

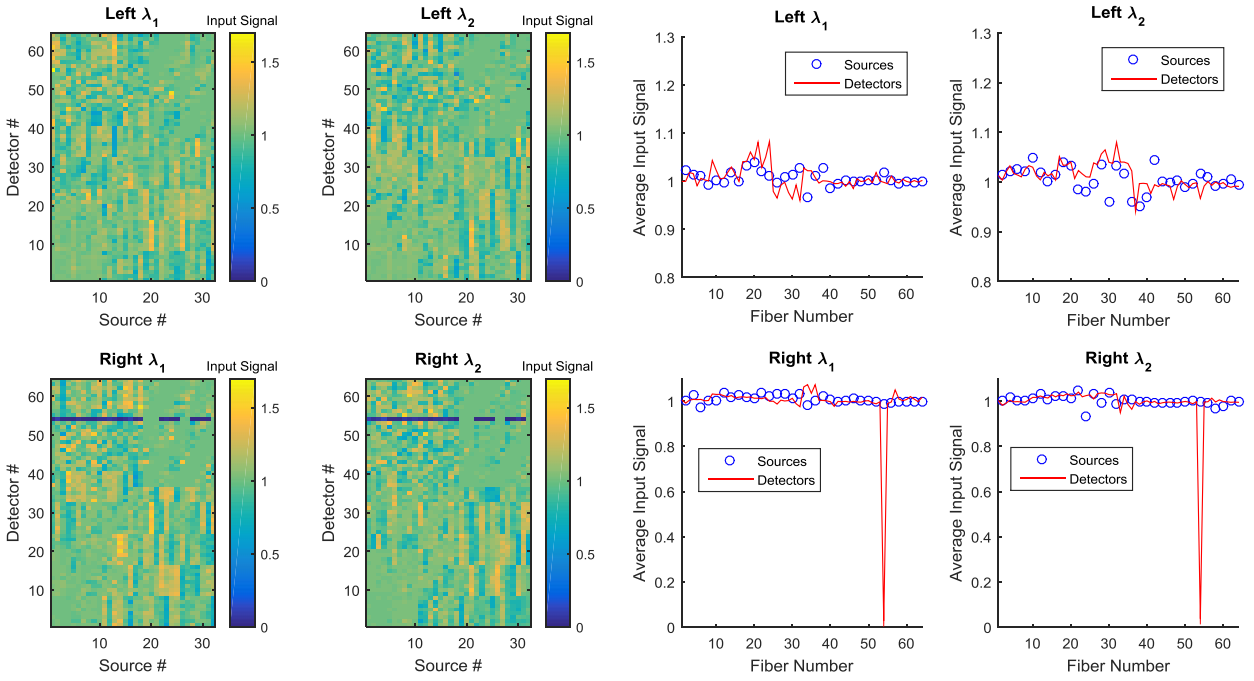


Figure 5.5 Visualization of input file data for each source detector pair signal for each wavelength for each breast (Left). The input signal is M_{tar}/M_{ref}. Also plots of the average input signal for each source and each detector (Right). Sources also act as detectors. The sources are represented by circles and overlay its corresponding detector signal. Detector 54 on the right side for both wavelengths shows a poor signal.

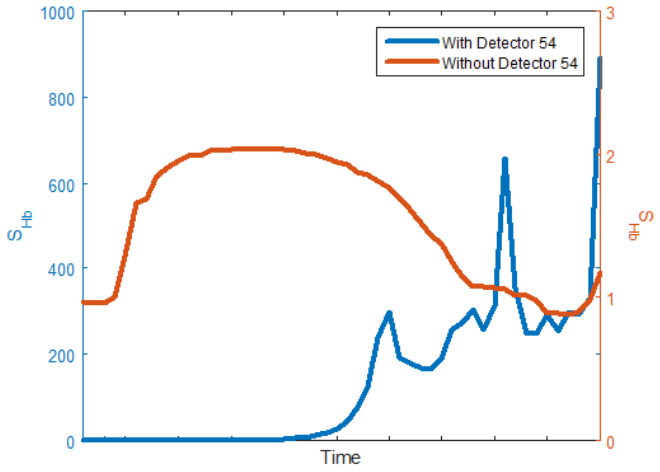


Figure 5.6 Time trace of the reconstructed results of the data seen in Figure 5.5 of the same breath hold with and without detector 54 removed.

Despite the current Matlab program having aided in increasing the speed and efficiency of pre-processing the dynamic data, there are still areas that could be improved. At the moment these poor source or detector signals are visualized and removed manually. However, the program could be improved if poor signals could be automatically removed.

5.4 Image Reconstruction Algorithm

The resulting data set was input to a diffusion-theory based PDE-constrained multispectral imaging algorithm.[69] This algorithm produced 3-dimensional volumetric reconstructions of various tissue parameters of both breasts. In this study we focused on the percent change of oxy-hemoglobin concentration (%Δ[HbO_2]) and percent change of deoxy-hemoglobin concentration (%Δ[Hb]) with respect to values observed at baseline. The baseline values were obtained as the average signal level over a period of 30 seconds prior to the patient's breath hold. Each frame was reconstructed using the

previous frame as the initial guess to the reconstruction, which decreases the reconstruction time and improves accuracy. The first frame starts with a uniform initial condition.

5.5 Post-Processing: Image Analysis and Quantification

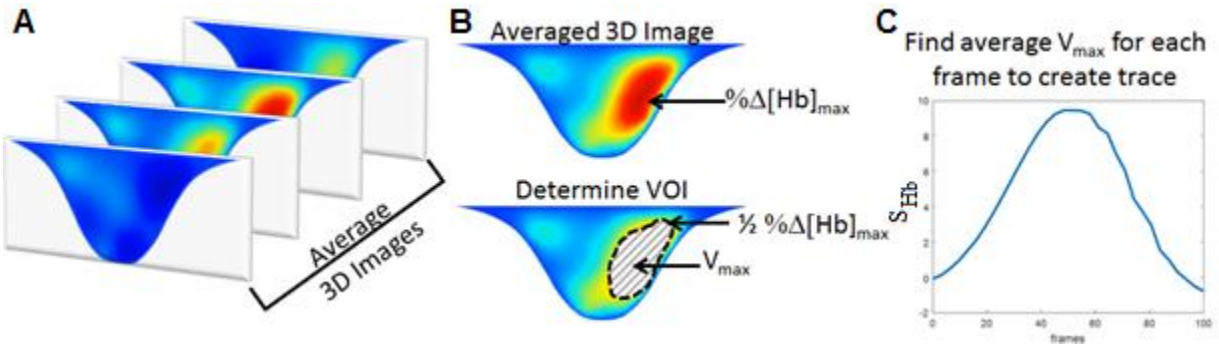


Figure 5.7 A visual of how time traces are acquired from 3D imaging. First the frames of the 3D images are averaged (A). The max of the averaged 3D image is found and a volume of half of the max is determined to become the volume of interest (VOI) (B). Last, the average of the voxels in the VOI for each frame is found and plotted (C).

After the reconstructions are completed, the movies are quantified. To identify the area that contains the tumor in the tumor-bearing breast, we average each voxel of the full 3-D tomographic data set over the period of the breath hold and half the recovery time. This results in a single frame of 3D data (see Fig. 5.7). Within this single frame we determine the location $\vec{x}_{max} = (x_{max}, y_{max}, z_{max})$ of the maximum percentage change in deoxyhemoglobin ($\%\Delta[Hb]_{max}$) and the value of half this change, defined as FWHM = $\%\Delta[Hb]_{max}/2$. All voxels within the vicinity of $\vec{x}_{max}$ for which $\%\Delta[Hb] \geq$ FWHM are defined as part of the "volume of interest" (VOI) which contains the region with the tumor. Then for each time point *t* of the original breath-hold data, the %Δ[Hb] values of each voxel within the VOI is averaged to get a single data point $S_{Hb}(t) = \%\Delta[Hb]_{avg}(t)$. This value will change from

time point to time point and a time-dependent curve $S_{Hb}(t)$ can be plotted (see e.g. Fig. 5.7). Similarly, the same process was used to determine $S_{HbO_2}(t) = \%\Delta[HbO_2]_{avg}(t)$.

5.6 Breath Hold Trace Improvement

Figure 5.8 shows an example of a breath hold for a patient before and after the above techniques were established. The breath hold in the before time trace (Fig. 5.8A) is difficult to distinguish and shows high frequencies. Figure 5.8B shows a much smother signal with a distinguishable breath hold. This improvement was due to the use of the DCT filter and the current post-processing methods that keep the VOI fixed for each frame.

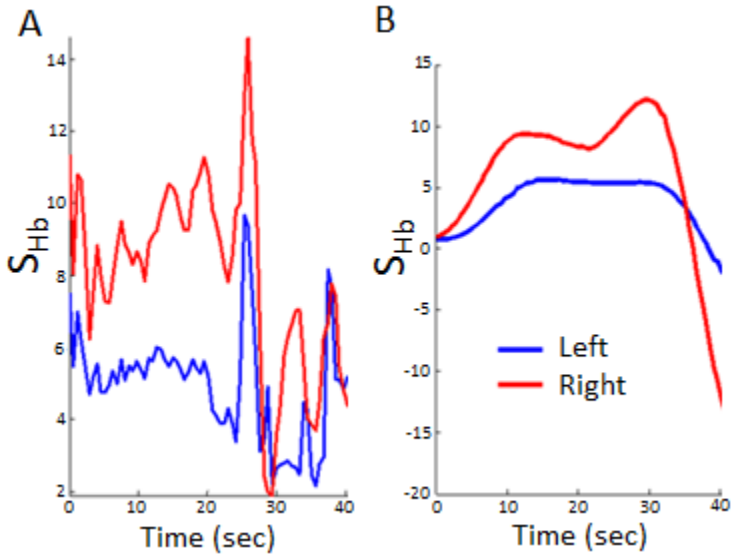


Figure 5.8 Example of a dynamic reconstruction of a baseline breath hold for a partial response patient using just a 10dB filter (A) and both a 10dB and DCT filter using the first 15 coefficients (B).

5.7 Feature Extraction

With reconstructed 3D imaging data sets of the two parameters (%Δ[Hb] and %Δ[HbO2]) for both breast in place, the task becomes to find suitable features in these images that show pronounced differences between the responders and non-responders.

First, a MatLab program was created to automatically find key points in the breath hold curve. Figure 5.9 shows the key points that were determined via this program.

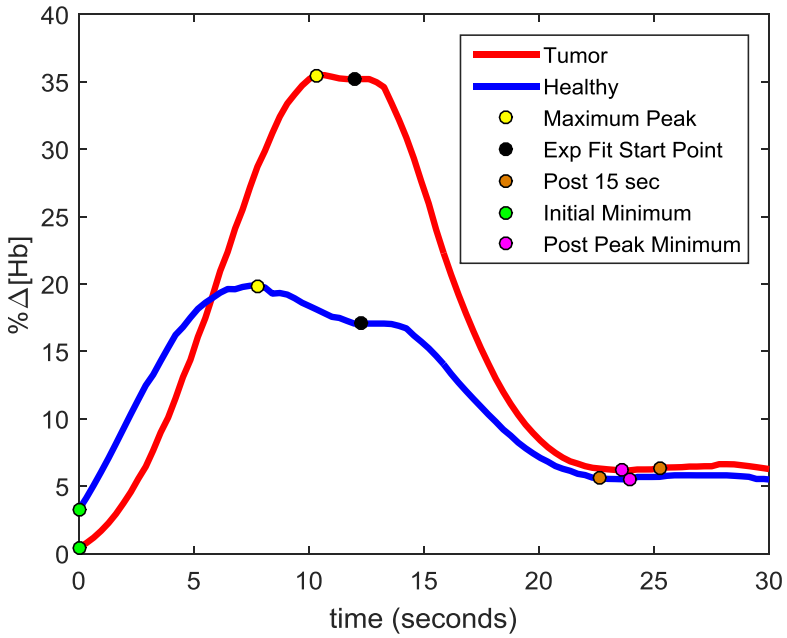


Figure 5.9 An example of a breath hold for both the tumor (red) and healthy (blue) undergoing feature extraction. Specific points along the breath hold curve were determined on the curves for further calculations. The maximum of the peak, the start point for the exponential fit for the recovery portion, the post 15 second after the peak, the initial minimum before the peak value, and the post-peak minimum value and location were determined.

The maximum peak is found by locating the maximum value during the breath hold. The 15 second post breath hold value is determined 15 seconds after the peak. To find the key minimum values in the graph, the local minimum before (initial) and after (post peak minimum) the peak were located. Last, the exponential fit start point is the maximum value before the recovery portion of the curve. This point is usually the peak value,

however, occasionally the breath hold would have a plateau section around the peak location (see Figure 5.9). Therefore, to properly determine the exponential fit the start point has to occur at the edge of this plateau portion of the breath hold. After these points were determined, the following features were determined.

Since time-dependent dynamic data was obtained during a breath hold, we first focused on time varying features inside the tumor. Figure 5.10 shows a plot of the time-dependent S_{Hb} values for the VOI in one of our subjects. One can see how the S_{Hb} increases sharply before a maximum is reached. Once the patient starts breathing again S_{Hb} decreases. To capture the features of this rise and fall in a more quantitative way, we introduce the following parameters, many of which have been used in other dynamic imaging modalities before: initial enhancement (IE) of the breath hold, post-initial enhancement (PIE), normalized maximum value (NMV), rise slope (m_{rise}), fall slope (m_{fall}), rise rate (τ_{rise}), and washout rate (τ_{fall}).

Initial enhancement (IE) is a parameter often employed in functional magnetic resonance imaging (MRI). It is a measure of how an MRI contrast agent increases in signal intensity after injections.[95] We adapted this parameter for DDOT by defining it as the ratio of the difference between the peak signal strength, S_{peak}, and the initial minimal signal before the peak $S_{initial,}$ divided the peak signal strength:

$$IE = \frac{S_{peak} - S_{initial}}{S_{peak}} \quad (1)$$

The ***post-initial enhancement (PIE)*** is another parameter used in contrast-enhance MRI. It is employed to determine how fast a contrast agent leaves a region an interest in DCE-MRI imaging. [95] For DDOT we define the PIE as

$$PIE = \frac{S_{postBH\ 15\ sec} - S_{peak}}{S_{peak}} \quad (2)$$

In this equation $S_{postBH15sec}$ is the [HbT] value 15 seconds post breath hold mark. In previous studies we found that this value leads to biggest differences between non-tumor bearing and tumor bearing breast.[23] The post-15 value was determined by using the value in the trace 15 seconds after the peak and normalizing it by the value of the peak.

The ***rise slope, m_{rise},*** was found by using the formula for slope between two points, where the first point was the minimum value before the peak and the second point was the peak point. The ***fall* slope, m_{fall},** was similarly found using the peak as the first point and the post-15 second point as the second. A visualization of the above features can be seen in Figure 5.10.

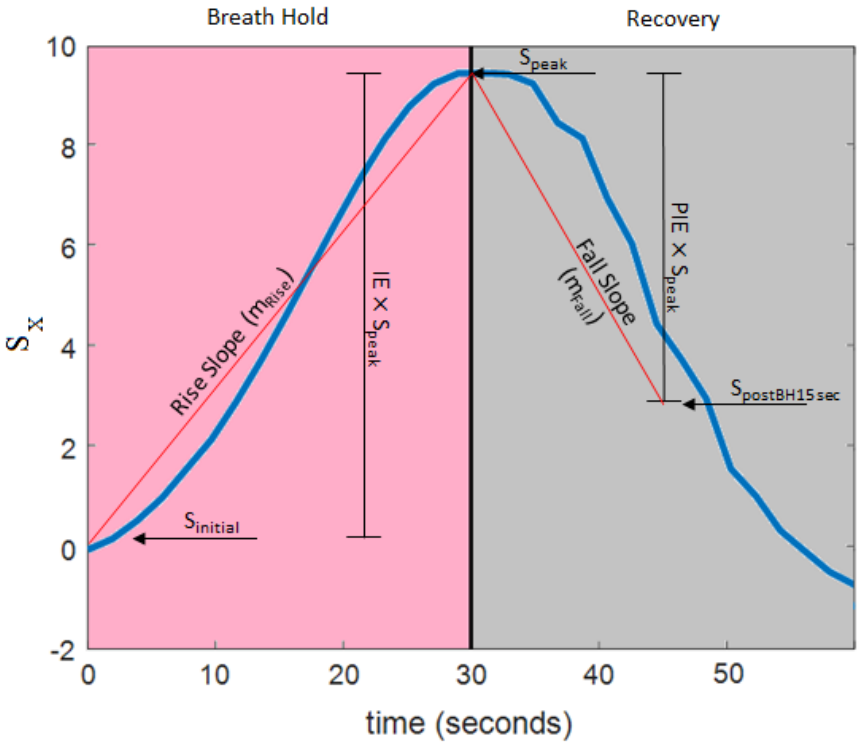


Figure 5.10 An example how several features are taken from the time trace for a single breast.

The ***rise rate, τ_{rise},*** was determined by fitting the rising portion of the curve to an exponential and taking the rate coefficient by looking at the section between the minimum

values before the peak to the peak value. Similarly, the ***washout rate, τ_{fall},*** was determined by fitting the recovery portion of the breath hold trace to an exponential and taking the rate coefficient. The recovery portion was defined from the maximum point of the trace to the following minimum position. These portions of the curve were fitted to the following equations:

$$S_X = A_{rise} * e^{\tau_{rise} * t}$$
$$S_X = A_{fall} * e^{-\tau_{fall} * t} \quad (3)$$

Where A_{rise} and A_{fall} were coefficients, *t* was the time that has passed in seconds, and S_X is either S_{Hb} or S_{HbO_2}. The parameters τ_{rise} and τ_{fall} were then used in the current analysis.

The ***normalized maximum peak value (NMPV),*** was calculated using the peak value of the tumor trace normalized by the peak value of the healthy breath hold trace.

In addition we calculated the ***correlation coefficient (CC)*** between the time traces observed in the healthy and tumor bearing breasts between the corresponding VOI.

In addition we tried various other parameters and features, such as time to maximum, area under the curve, etc. However, these other features did not show promising results, and hence we focused on the 9 parameters just introduced here.

5.8 Conclusion

By introducing the DCT filter and the automated program, dynamic DOT data was able to be processed quickly and efficiently. Also, by developing methods to visualize the data there has been progress to create check points that assist in breath hold choice and quality. Therefore, the methods explained above has greatly contributed to the quality of

the dynamic breath hold time traces. Last, features were created and calculated to assist in the data analysis to determine if the hemodynamic nature of the breath hold can be used to determine tumor response to NACT.

Chapter 6: Dynamic DOT Results

6.1. Introduction

In addition to the static images that we acquired using our DOT system, we had acquired and reconstructed movies of the hemodynamics of subjects holding their breath. This is the largest dynamic NACT DOT monitoring study at the moment. Other groups differ in imaging modality and methodology, however, few groups have employed dynamic techniques for the purpose of observing hemodynamic changes in the NACT setting.

As previously mentioned, given the non-invasive character of optical imaging modalities and its sensitivity to blood-dependent parameters and tissue vascularity, several groups have evaluated DOT or non-tomographic diffuse optical spectroscopy (DOS) for monitoring response to NACT. For example, Ueda et. al. correlated the pCR rates to baseline tumor oxygen saturation measurements obtained with a handheld spectroscopic measurement system.[60] The handheld probe had only one source and one detector that were scanned across a region of interest to take both frequency-domain and steady-state measurements. No optical tomographic reconstruction of the entire breast was performed. Other groups employed tomographic breast imaging systems.[36, 57, 74, 80, 96] This includes the Zhu group at the University of Connecticut, who performed ultrasound guided DOT NACT monitoring studies;[58, 72, 73, 78, 79] and Dr. Yodh's team at UPENN, who performed a pilot study in 3 patients undergoing NACT. They employed a combined frequency and continuous-wave system that uses two parallel plates to hold one of the breasts in place inside a "breast box." The box is filled with a solution of Intralipid and India ink that is used as matching fluid and reference medium.[36, 96] These groups have

either employed dual-modality OI or DOS to perform NACT monitoring in which chromophore concentrations were observed, but have not utilized any dynamic capabilities to detect hemodynamic trends of breast tumors.

In addition to the previous mentioned aspects in Chapter 5, the study differs from previous research in a very distinct way. We employ a full-tomographic 3-dimensional dynamic DOT imaging that provides information about hemodynamic effects of a breath hold in the NACT setting.

Our group has previously used a dynamic diffuse optical tomographic (DDOT) system to observe the hemodynamics of different tumors to determine if a subject had a benign or malignant tumor or no tumor at all. The contrast mechanism that was used for that diagnostic study was a simple breath hold. The quantified hemodynamic time traces showed variation between the healthy and tumor bearing breasts of subjects compared to those of healthy subjects, which showed similar traces for both breast. Additionally, we observed a more pronounced change in deoxy-hemoglobin (Δ[Hb]%) in the tumor at the 15 second post breath hold.[23]

Given these promising results, we hypothesized that we will be able to observe the vascular changes that occur in tumors utilizing DDOT early in treatment. In particular, we expected that information about early changes in the hemodynamics can be used to predict the NACT response of individual patients. To test this hypothesis, we designed a longitudinal clinical study that uses DDOT to observe tumor response to throughout 5-months of NACT.

6.2. Results

6.2.1 Patient Summary

There were forty women recruited and imaged for the study with four subjects being excluded due to drop outs or machine malfunctions. Table 2.2 show a subject summary of the patients and their pathological response.

Since for this chapter we are only concerned with the baseline and week 2 time points. Results from 6 patients were not included in the analysis presented here. One patient completely stopped NACT midway through the therapy and was removed from the study. For another 2 subjects, instrumentation failure during the imaging session resulted in corrupted data sets that could not be used. One subject opted to receive radiation therapy in addition to the traditional NACT therapy and therefore was removed from the analysis. Two subjects completely missed there week 2 imaging time point.

6.2.2 Case Studies

Before a full statistical analysis is provided, we present two case studies that illustrate the overall results we obtained. The first case involves a 46 year-old post-menopausal woman with a poorly differentiated invasive ductal carcinoma. Her baseline tumor, located in the left breast, had an initial maximum diameter of 2.9 cm as determined by mammography (a clinical exam found the diameter to be 5.5 cm). The patient had a pCR to NACT (RCB-0). Figure 6.1A shows the baseline mammogram of her left breast with calcifications in the outer quadrants. Baseline optical images shows a greater percent change in [Hb] than in the healthy breast at the 15 second post-breath hold time frame (Fig. 6.1B). Figure 6.1C shows the time traces of changes in deoxyhemoglobin (S_{Hb}) in response to a breath hold obtained at baseline and 2 weeks after the start of NACT. We observe that while the rise slopes of both curves are similar, the washout rate

at the two-week time point (black curve) is noticeably higher compares to baseline (red curve).

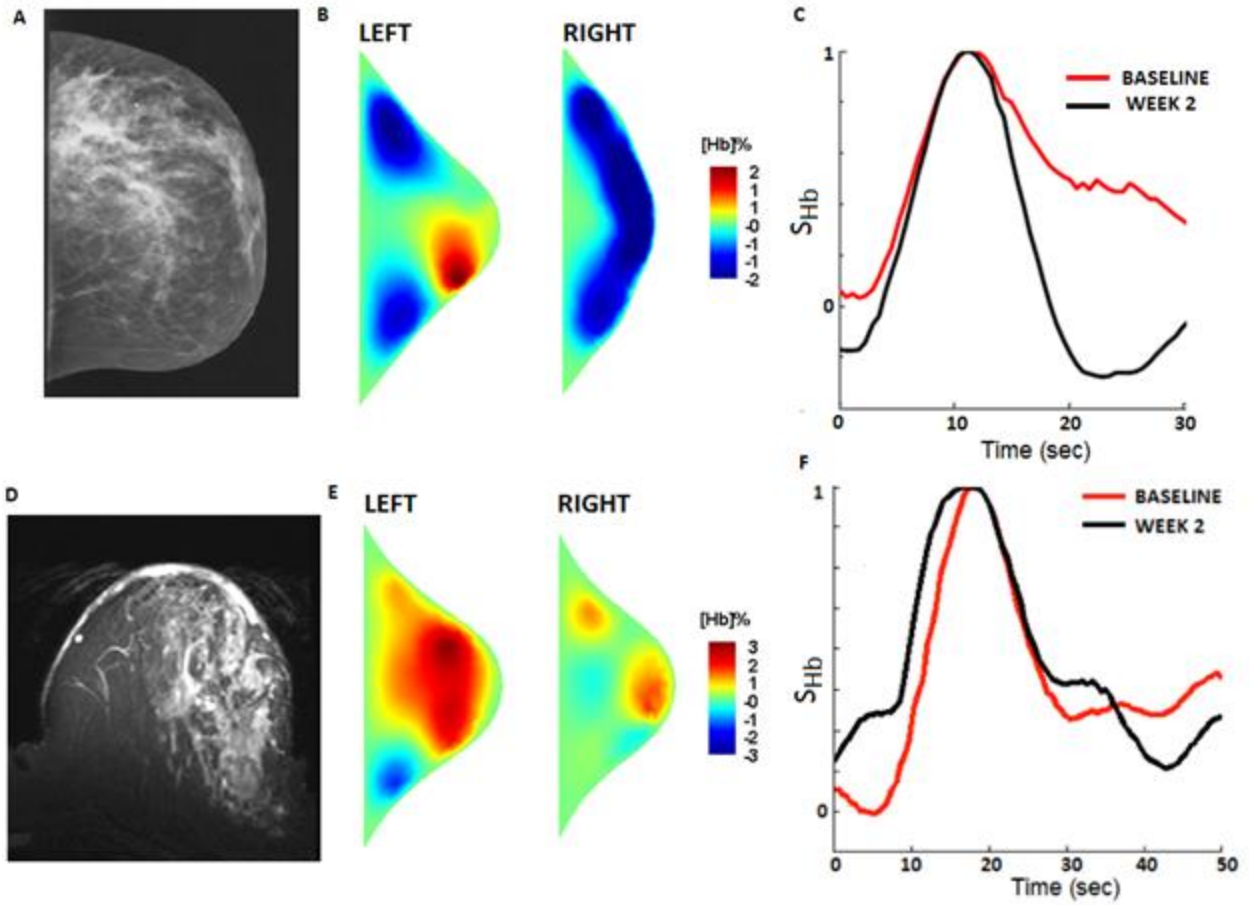


Figure 6.1 The top row shows imaging results from a patient with a pathological complete response (RCB-0); the bottom row shows imaging results from a patient with no response (RCB-III) to NACT. Figures A and D display magnetic resonance images (MRIs) obtained prior to chemotherapy. Figures B and E are DOT images obtained from the left and right breasts just before therapy. The images displayed refer to a time point of 15 seconds after the breath hold; Finally, in figures C and F we show time-dependent signal traces in response to a 30 second breath hold for the tumor region obtained from DDOT just before the start of NACT (baseline) and 2 weeks after treatment initiation.

The second case study involves a 61 year-old post-menopausal woman with a moderately differentiated invasive lobular carcinoma. Employing ultrasound imaging and x-ray mammography it was determined that she had a very large tumor of 9.5 cm in

diameter in the left breast, at the start of the therapy. She did not respond well to NACT and was classified as a RCB-III patient. The patient's baseline DOT images at the 15 second post-breath hold time frame show a large enhancement in the left tumor-bearing breast, while the healthy breast shows a very small increase in the percent change of [Hb] (Figure 6.1E). Furthermore, there is little difference between the time-trace of S_{Hb} obtained at baseline and 2-weeks into chemotherapy (Fig. 6.1F). Especially, the washout rates are almost the same, which is in stark difference to the time-traces observed in the patient with a pCR (Fig. 6.1C).

6.2.3 Statistical Analysis of Hemodynamic Parameters

6.2.3.1 Analysis of Hemoglobin Traces observed at Week 2

We start our analysis by looking at the hemodynamic traces observed 2 weeks after the first treatment with taxane. A summary of our main findings is given in Table 6.1. The table shows the means values for all four cohorts (RCB-0, I, II, III) for all 9 features (IE, PIE, m_{Rise}, m_{Fall}, τ_{rise}, τ_{fall}, S_{Peak}, NMPV, and CC) of both the S_{Hb} and S_{HbO_2} response curves. Furthermore the values for various ANOVA analyses and p-values are given. Statistical significant differences are highlighted in green ($0.01 \leq p \leq 0.05$) and yellow ($p < 0.01$). The m_{rise}, S_{peak}, NMPV, and CC do not show any statistically significant difference among or between any of the groups for both S_{Hb} and S_{HbO_2} traces. The features that show some statistically significant differences between groups are IE, PIE, m_{fall}, τ_{rise}, and τ_{fall}. As these features promise to be good predictors for treatment outcome, we will discuss each them in more detail in the following sections.

Table 6.1 Summary of features derived from time traces of the hemodynamic responses (both S_{Hb} and S_{HbO_2}) obtained 2 weeks after the initial treatment with taxane. Shown are the mean and standard deviation for 9 different hemodynamic-curve features (IE, PIE, m_{rise}, m_{fall}, τ_{rise}, τ_{fall}, Speak, NMPV, and CC) for all four sub-cohorts (RCB 0, I, II, and III). One way ANOVA's were performed among all four groups, then three groups with RCB I subjects grouped with either RCB 0 or RCB II subjects. Fields highlighted in green mark statistical significance ($p<0.05$); yellow denotes "strong" statistical significance ($p<0.01$), and blue fields have values of $0.05<p<0.1$.

Features	Time - Traces	RCB 0	RCB I	RCB II	RCB III	ANOVA 1 (4 groups)	ANOVA 2 (3 groups 0&I)	ANOVA 3 (3 groups I&II)	pCR vs non-pCR (RCB 0 vs RCB I, II & III)	RCB 0&I vs RCB II & III	RCB 0,I,&II vs RCB III
IE	S_{Hb}	1.02 ± 0.28	0.92 ± 0.04	0.96 ± 0.15	0.71 ± 0.34	0.09	0.04	0.04	0.13	0.15	0.01
	S_{HbO2}	1.03 ± 0.29	0.90 ± 0.03	0.96 ± 0.18	0.73 ± 0.35	0.14	0.08	0.07	0.13	0.17	0.03
PIE	S_{Hb}	-0.89 ± 0.52	-0.51 ± 0.18	-0.40 ± 0.42	-0.30 ± 0.56	0.04	0.02	0.01	0.004	0.01	0.15
	S_{HbO2}	-0.77 ± 0.38	-0.47 ± 0.21	-0.49 ± 0.46	-0.44 ± 0.48	0.30	0.24	0.15	0.052	0.09	0.37
m_{rise}	S_{Hb}	0.80 ± 0.83	0.47 ± 0.21	0.70 ± 0.49	0.38 ± 0.46	0.58	0.46	0.41	0.34	0.47	0.22
	S_{HbO2}	0.44 ± 0.47	0.24 ± 0.06	0.32 ± 0.20	0.17 ± 0.15	0.39	0.30	0.23	0.14	0.21	0.17
m_{fall}	S_{Hb}	-0.46 ± 0.32	-0.41 ± 0.31	-0.19 ± 0.20	-0.01 ± 0.34	0.015	0.005	0.008	0.01	0.002	0.03
	S_{HbO2}	-0.24 ± 0.25	-0.19 ± 0.13	-0.11 ± 0.12	-0.03 ± 0.14	0.11	0.05	0.054	0.03	0.02	0.09
τ_{rise}	S_{Hb}	0.11 ± 0.06	0.07 ± 0.004	0.17 ± 0.10	0.05 ± 0.04	0.03	0.01	0.04	0.70	0.43	0.03
	S_{HbO2}	0.12 ± 0.06	0.07 ± 0.01	0.17 ± 0.10	0.08 ± 0.06	0.10	0.06	0.17	0.61	0.32	0.13
τ_{fall}	S_{Hb}	0.18 ± 0.09	0.07 ± 0.02	0.09 ± 0.10	0.07 ± 0.07	0.03	0.05	0.01	0.003	0.02	0.16
	S_{HbO2}	0.17 ± 0.14	0.07 ± 0.04	0.11 ± 0.10	0.11 ± 0.07	0.38	0.44	0.23	0.09	0.20	0.59
S_{peak}	S_{Hb}	9.24 ± 7.54	11.0 ± 5.21	8.61 ± 5.31	9.26 ± 8.31	0.97	0.94	0.99	0.93	0.77	0.95
	S_{HbO2}	5.34 ± 5.29	5.62 ± 1.66	4.34 ± 3.29	3.98 ± 3.82	0.88	0.72	0.78	0.50	0.42	0.63
NMPV	S_{Hb}	1.83 ± 1.62	1.16 ± 0.85	1.55 ± 1.24	2.93 ± 4.27	0.61	0.43	0.41	0.90	0.72	0.20
	S_{HbO2}	1.85 ± 1.51	0.81 ± 0.26	1.32 ± 0.95	2.57 ± 3.39	0.48	0.39	0.31	0.75	0.97	0.20
CC	S_{Hb}	0.46 ± 0.38	0.57 ± 0.17	0.36 ± 0.34	0.43 ± 0.36	0.34	0.52	0.96	0.77	0.28	0.93
	S_{HbO2}	0.53 ± 0.37	0.69 ± 0.26	0.41 ± 0.34	0.39 ± 0.37	0.47	0.42	0.69	0.45	0.19	0.50

Since there are four response groups, there are several ways to create groupings of the subjects. Most studies try to find the patients with a pCR (RCB-0) and group patients in classes RCB-I, -II, and -III into one class of non-pCRs. However, if the goal is to find patients who would benefit from a change in therapy, therefore patients that do not respond to NACT, one could argue that finding patients that belong to the RCB-III class

is equally important. From our statistical analysis, it is apparent that *IE* and τ_{rise} are the two best features for finding RCB III subjects. The p-values for differences between RCB-III and all other patients (last column in Table 6.1) are 0.01 and 0.03 for IE and τ_{rise} determined for the S_{Hb} traces. As seen in Figure 6.2A, subjects that had some response to NACT (RCB 0, I, &II) have a higher IE value at week 2 compared to RCB III subjects for both the S_{Hb} and S_{HbO_2} traces. Similarly, the responding subjects have a higher τ_{rise}, but only the S_{Hb} traces showed a significant difference between response groups (Figure 6.2B).

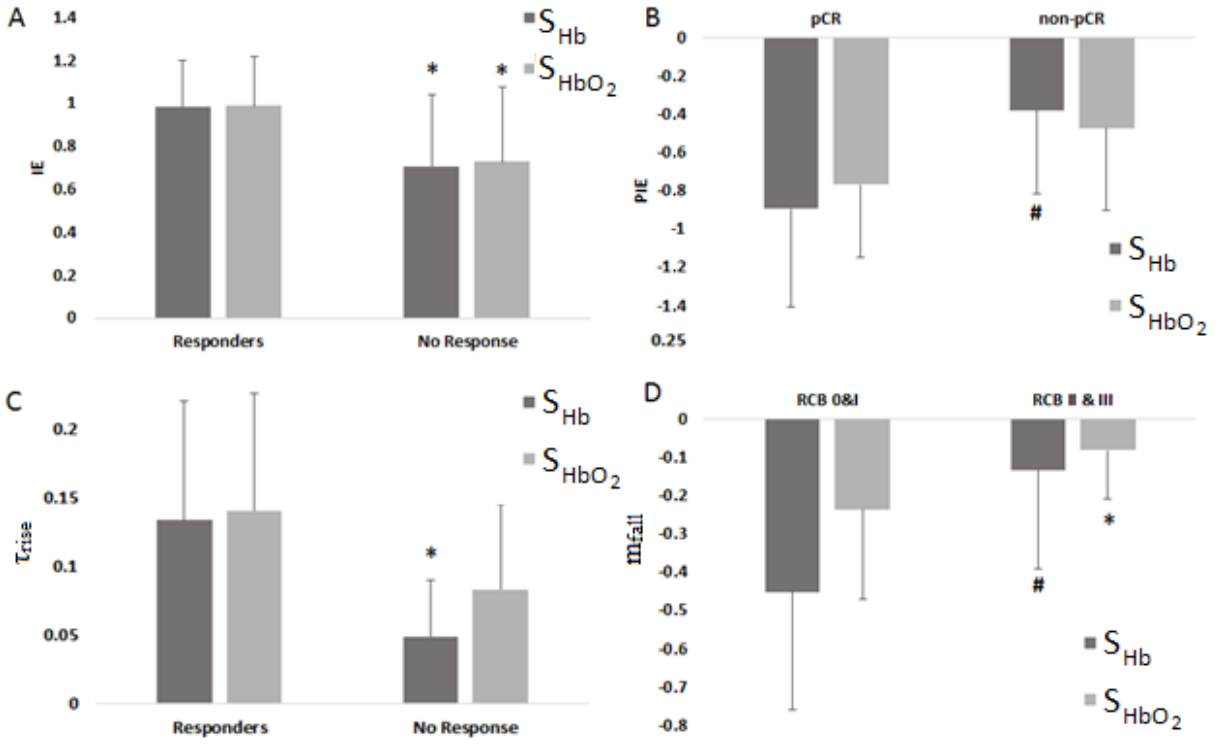


Figure 6.2 Comparison of different groupings for statistically significant features at week 2. Week 2 IE value for subject groups responders versus no response (A). Week 2 PIE value for pCR and non-pCR subject groups (B). Week 2 τ_{rise} coefficient for subjects that had some response (RCB 0, I, & II) compared to no response subjects with an RCB of III (C). The m_{fall} at week 2 for RCB 0&I and RCB II&III groups (D). Statistical significance between the two response groups is denoted by (*) for $p<0.05$, (#) for $p<0.01$, and (&) for $p<0.001$.

If one wanted to identify RCB-II and RCB-III group members as candidates who benefit from a change in treatment then the RCB 0&I versus RCB II&III grouping should be used (see next to last column in Table 6.1). For this case we observed the largest statistical significance ($p<0.002$) in the m_{fall} for the S_{Hb} traces. The RCB 0&I subjects showed steeper slopes at week 2 compared to RCB II &III subjects for both the S_{Hb} and S_{HbO_2} traces.

Similar to other studies we also looked for parameters and features for which pCR subject could be distinguished best from non-pCR subjects (RCB-0 vs RCB-I, -II, -III). In this case the most significant parameters were PIE ($p<0.004$) and τ_{fall} ($p<0.003$). The pCR subjects had lower PIE values at week 2 for both the S_{Hb} ($p<0.004$) and S_{HbO_2} ($p<0.052$) traces compared to non-pCR subjects (Figure 6.1D). Also, pCR subjects had faster τ_{fall} compared to non-pCR subject, but only S_{Hb} traces showed statistical significance ($p<0.03$) (Figure 6.3).

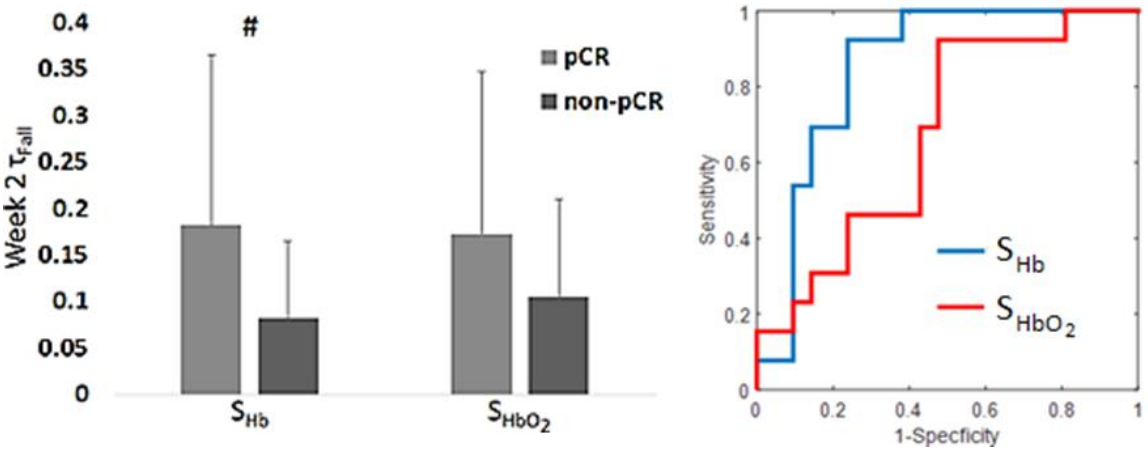


Figure 6.3 Week 2 τ_{fall} coefficient for subjects that had pCR compared to non-pCR. There is only statistical significance for the S_{Hb} parameter with (#) p=0.003 (%d [HBO2]: p=0.08) (Left). ROC curve for both S_{Hb} and S_{HbO_2} with an area under the curve of 85 and 67, correspondingly (Right).

6.2.3.2 ROC Curves for selected Week 2 features

Since the difference in m_{fall} for S_{Hb} traces between patients with a pCR and non-pCR produced the strongest statistically significant distinction, we created a ROC curve for this case (see Figure 6.3). The sensitivity, specificity, and accuracy for the washout rate for S_{Hb} was 92.3% (12/13 subjects), 76% (16/21 subjects), and 82.4% (28/34 subjects), respectively (cut-off point τ_{fall} = 0.09). The area under the curve (AUC) was 0.85, and the PPV calculated to 70.6% with the NPV 94.1%. A similar ROC analysis for m_{fall} and PIE for S_{Hb} resulted in values shown in Table 6.2.

As for differentiating the RCB-0&I subjects from RCB-II&III the parameter with smallest p-value for this analysis was m_{fall} for S_{Hb}. The corresponding ROC curve has an AUC of 0.83 with a sensitivity of 86.7% (13/15 subjects), and specificity of 73.7% (14/19 subjects), and an accuracy of 79.4% (27/34 subjects). The PPV was 72.2% and the NPV was 85.7%. Additionally, m_{fall} for S_{HbO_2} was statistically significant (p = 0.02) and the ROC analysis can be found in Table 6.2.

Table 6.2 Summary of binary classification test for the three features that show the statistically most significant differences between patients with a pCR (RCB-0) and non-pCR patients (RCB-I, -II, -II), based on an analysis of the data obtained at week 2. Sensitivity is the percent of the pCR subjects that were classified as pCR. Similarly, specificity is the percent of the non-pCR subject that were classified as non-pCR.

pCR v non-pCR	Sensitivity	Specificity	Accuracy	AUC	PPV	NPV
m_{fall}	92.3%	76.0%	82.4%	0.85	70.6%	94.1%
PIE	53.8%	90.5%	76.5%	0.73	77.8%	76.0%
RCB 0&I v. II&III	Sensitivity	Specificity	Accuracy	AUC	PPV	NPV
m_{fall} for S_{Hb}	86.7%	73.7%	79.4%	0.83	72.2%	87.5%
m_{fall} for S_{HbO_2}	86.7%	63.2%	73.5%	0.77	65.0%	85.7%
RCB 0,I&II v. III	Sensitivity	Specificity	Accuracy	AUC	PPV	NPV
IE for S_{Hb}	96.4%	50.0%	88.2%	0.71	90.0%	75.0%
IE for S_{HbO_2}	96.4%	50.0%	88.2%	0.66	90.0%	75.0%
τ_{rise} for S_{Hb}	53.6%	100%	61.8%	0.83	100.0%	31.6%

For differentiation between RCB-III and all other patients, the parameter IE determined from the S_{Hb} traces produced the smallest p-value (p=0.01). The

corresponding ROC curve has an AUC of 0.71 with an accuracy of 73.5% (30/34 subjects), sensitivity of 96.4% (27/28 subjects), and specificity of 50.0% (3/6 subjects). The PPV is 90.0% and the NPV is 75.0%. The sensitivity represents the ability to predict responding subjects (RCB-0, -I, and –II) from the non-responders. The sensitivity, specificity, accuracy, AUC, PPV, and NPV for the other statistically significant parameters can be seen in Table 6.2.

6.2.4 Analysis of Differences between Baseline and Week 2 Imaging

In the previous section we only looked at data obtained two weeks after initiation of the NACT protocol. To determine if additional parameters and features could be found that yield even better separation between various RCB classes, we also analyzed ***differences*** between data obtained at baseline (week 0) and week 2. Table 6.3 shows a summary of each of the features.

In this table, the most notable entries are the low p-values for Δm_{fall} (changes in the fall slope) determined from S_{Hb} traces (highlighted in yellow). Of the 6 different groupings considered, only RCB-III vs RCB-0, -I, -II, shows a non-significant statistical difference ($p=0.64$). All other groupings show p-values ≤ 0.001. The features Δm_{rise} and Δm_{fall} had the lowest p-values for pCR versus non-pCR and can be seen in Figure 6.4. All features show an increase for pCR subjects and a decrease for non-pCR subjects. For the particularly important subgrouping of pCR versus non-pCR we find that the sensitivity, specificity, and accuracy is 61.5% (8/13 subjects), 100% (21/21 subjects), and 85.3% (29/34 subjects) with an AUC of 0.87. The PPV was 100% and the NPV was 80.8%. The 100% specificity signifies that all the patients with a non-pCR were identified as having a non-pCR response. This is of particular importance, as these are exactly the

patients that would benefit form a change in therapy. The ROC curves for the significant parameters for this grouping can be found in Figure 6.5A&B.

Table 6.3 Summary of the differences in features between baseline and two weeks after the initial treatment with taxane (Week 2 – Baseline). Shown are the mean and standard deviation for 9 different hemodynamic-curve features (IE, PIE, m_{rise}, m_{fall}, τ_{rise}, τ_{fall}, Speak, NMPV, and CC) for all four sub-cohorts (RCB 0, I, II, and III). One way ANOVA's were performed among all four groups, then three groups with RCB I subjects grouped with either RCB 0 or RCB II subjects. Green marks statistical significance ($0.01 \leq p \leq 0.05$). Yellow denotes "strong" statistical significance ($p<0.01$), and blue shows entries with $0.05<p\leq0.1$).

Features	time traces	RCB 0 (pCR)	RCB I	RCB II	RCB III	ANOVA 1 (4 groups)	ANOVA 2 (3 groups 0&I)	ANOVA 3 (3 groups I&II)	pCR vs non-pCR (RCB 0 vs RCB I, II & III)	RCB 0&I vs RCB II & III	RCB 0,I,&II vs RCB III
ΔIE	S_{Hb}	0.03 ± 0.39	-0.39 ± 0.27	-0.04 ± 0.30	-0.29 ± 0.31	0.16	0.28	0.19	0.17	0.47	0.11
	S_{HbO2}	0.16 ± 0.38	-0.35 ± 0.49	0.03 ± 0.41	-0.40 ± 0.37	0.03	0.05	0.02	0.055	0.18	0.01
ΔPIE	S_{Hb}	0.29 ± 0.66	0.15 ± 0.45	-0.25 ± 0.70	0.02 ± 0.53	0.23	0.12	0.15	0.07	0.06	0.97
	S_{HbO2}	0.17 ± 1.08	0.08 ± 0.77	-0.25 ± 0.54	0.02 ± 0.64	0.63	0.42	0.48	0.28	0.26	0.87
Δm_{rise}	S_{Hb}	0.47 ± 0.65	-0.06 ± 0.01	-0.13 ± 1.08	-0.07 ± 0.26	0.27	0.20	0.14	0.04	0.07	0.55
	S_{HbO2}	0.26 ± 0.36	0.003 ± 0.001	-0.03 ± 0.51	-0.07 ± 0.14	0.22	0.15	0.11	0.03	0.052	0.34
Δm_{fall}	S_{Hb}	0.33 ± 0.33	0.03 ± 0.27	-0.38 ± 0.56	-0.13 ± 0.30	0.003	0.001	0.002	0.0004	0.0004	0.64
	S_{HbO2}	0.16 ± 0.21	-0.08 ± 0.31	-0.12 ± 0.27	-0.06 ± 0.13	0.03	0.03	0.01	0.0025	0.01	0.50
$\Delta \tau_{rise}$	S_{Hb}	-0.04 ± 0.14	-0.1 ± 0.21	-0.03 ± 0.19	-0.12 ± 0.22	0.64	0.55	0.60	0.69	0.94	0.31
	S_{HbO2}	0.002 ± 0.10	-0.13 ± 0.20	0.001 ± 0.18	-0.14 ± 0.12	0.15	0.14	0.13	0.32	0.59	0.05
$\Delta \tau_{fall}$	S_{Hb}	0.02 ± 0.14	-0.01 ± 0.05	-0.11 ± 0.19	-0.07 ± 0.12	0.23	0.11	0.16	0.056	0.04	0.70
	S_{HbO2}	0.01 ± 0.19	-0.01 ± 0.06	-0.14 ± 0.19	-0.06 ± 0.11	0.23	0.11	0.17	0.08	0.055	0.98
ΔS_{peak}	S_{Hb}	-4.44 ± 4.96	-1.32 ± 1.68	2.13 ± 11.6	-3.53 ± 8.58	0.14	0.16	0.56	0.14	0.16	0.56
	S_{HbO2}	-2.68 ± 3.20	-0.49 ± 2.36	-0.19 ± 6.23	-1.27 ± 4.30	0.20	0.26	0.96	0.20	0.26	0.96
ΔNMPV	S_{Hb}	-0.74 ± 1.83	0.15 ± 0.06	-0.02 ± 1.81	-2.19 ± 4.61	0.37	0.23	0.21	0.93	0.89	0.11
	S_{HbO2}	-0.63 ± 1.86	0.23 ± 0.23	0.41 ± 1.75	-1.92 ±3.67	0.22	0.12	0.11	0.70	0.86	0.08
ΔCC	S_{Hb}	0.01 ± 0.41	0.07 ± 0.18	0.22 ± 0.45	-0.04 ± 0.38	0.50	0.32	0.33	0.39	0.41	0.41
	S_{HbO2}	-0.02 ± 0.45	0.04 ± 0.27	0.07 ± 0.50	0.02 ± 0.33	0.93	0.92	0.83	0.58	0.73	0.98

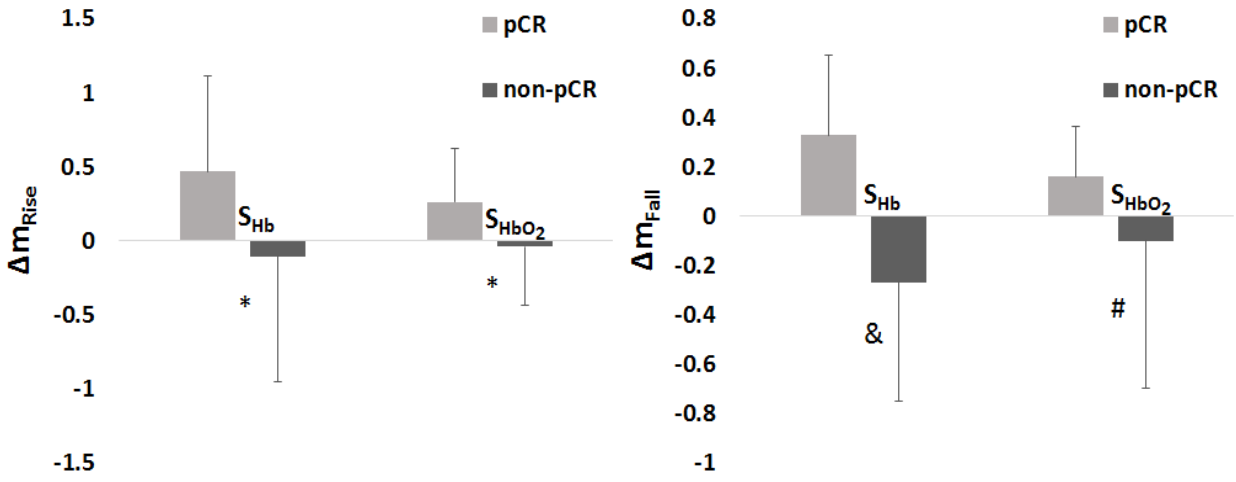


Figure 6.4 Comparison of pCR and non-pCR for Δm_{rise} (Left), Δm_{fall} (Right), between baseline and week 2 imaging. Statistical significance between the two response groups is denoted by (*) for p<0.05, (#) for p<0.01, and (&) for p<0.001.

To separate RCB 0&I subjects from the RCB II&III subjects, τ_{fall} had statistical significance for S_{Hb} traces in which RCB II&III showed decreases in washout rate and pCR subjects had an increase in washout rate. Figure 6.5C shows the ROC curve for $\Delta\tau_{fall}$ for S_{Hb}. The AUC was 0.76 and the overall accuracy at the Youden index was 76.5% (26/32 subjects). The corresponding sensitivity, specificity, PPV, and NPV was 73.3% (11/15), 78.9% (15/19), 71.3%, and 78.9%.

To sort out RCB-III patients, the best features seem to be IE and τ_{rise} for the S_{HbO_2} traces (p=0.01, and 0.05, respectively). The AUC for IE and τ_{rise} for the S_{HbO_2} traces was 0.83 and 0.76 (Figure 6.5D). Looking at the higher AUC, τ_{rise} for the S_{HbO_2} had a sensitivity, specificity, and accuracy of 85.7%, 83.3%, and 85.3%, respectively. Here, the sensitivity is the percentage of responders (RCB-0, -I, &-II) that were correctly identified. The PPV was 96.0% and a NPV of 55.6%.

PIE, S_{Peak}, NMVP, and CC did not show any statistical significance between or among the different response groups.

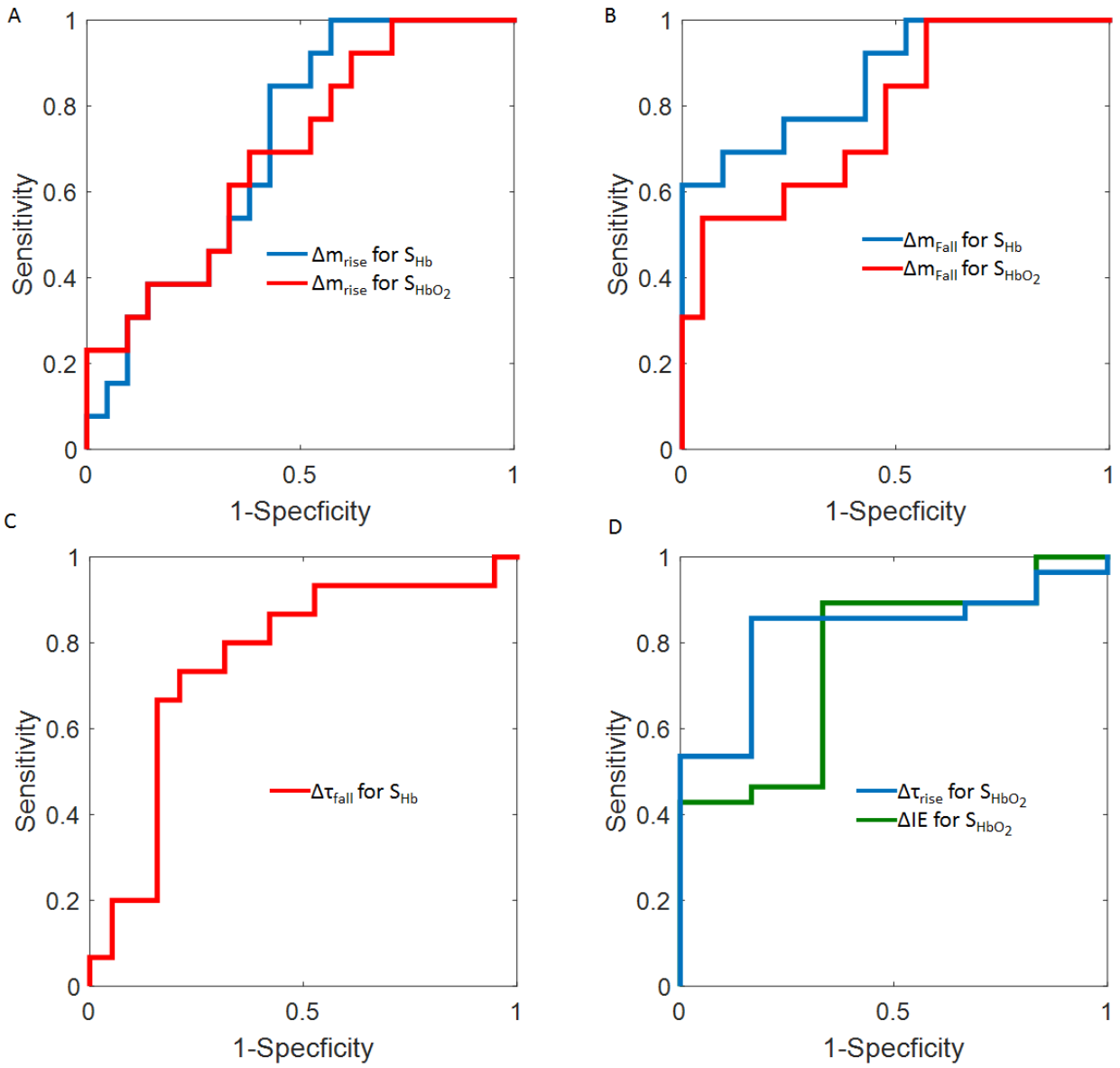


Figure 6.5 ROC curves for select parameters from Table 6.3 in which the RCB 0 were compared to RCB-I, -II, and –III subjects (A and B), the RCB 0&I were compared to RCB-II&III (C), and last where the RCB-0, -I, and –II were compared to RCB-III subject (D). The AUC for Δm_{rise} for S_{Hb} and S_{HbO_2} was 0.71 and 0.69, respectively. For Δm_{fall} the AUC for S_{Hb} was 0.87 and for S_{HbO_2} was 0.78. The ROC curve for $\Delta \tau_{fall}$ for S_{Hb} had an AUC of 0.76. The AUC for the $\Delta \tau_{rise}$ and ΔIE for S_{HbO_2} was 0.83 and 0.76, correspondingly.

6.3. Discussion

6.3.1 Summary of Results

We have performed a clinical pilot study to evaluate if DDOT can be used to predict pathologic response to NACT setting. In this study, 40 patients were enrolled, of which dataset from 6 had to be discarded, either because the patients drop out of the study before it was finished, or instrumentation malfunction. We evaluated a total of 9 curve parameters for time-dependent traces of S_{Hb} and S_{HbO2} recorded inside the tumor. Recordings were obtained at baseline, just before treatment initiation (week 0) and 2 weeks after NACT started. Our study showed that tumors in patients with pCR (RCB = 0) have higher washout rates, m_{fall}, by week 2 compared to non-pCR (RCB=I, II, III) patients. Using m_{fall} for S_{Hb} as parameter to predict patients with a pCR, we achieved a sensitivity of 92.3% and a specificity of 76%. Here sensitivity is defined as the ability to determine if the patient will have a pCR, and specificity is how well the non-pCR subjects can be determined. Additionally, the Δm_{rise} and Δm_{fall} showed significant difference between pCR and non-pCR subjects, where pCR subjects showed an increase in each parameter for both S_{Hb} and S_{HbO2} traces. This may possibly indicate and increase in blood flow in the tumor. For this analysis, the highest AUC (0.87) and specificity (100%) were obtained for Δm_{fall} (fall slope) of the S_{Hb} trace.

In comparing RCB-0/I from RCB-II/III, the best parameter would be m_{fall} at week 2 for S_{Hb} where an accuracy of 79.4%, the highest for this grouping. However, the best test for determining RCB III was using $\Delta\tau_{rise}$ for S_{HbO2} in which the corresponding ROC curve had the highest Youden index with a sensitivity of 85.7% and a specificity of 83.3%.

IE and PIE show very similar trends to the τ_{rise} and τ_{fall} parameters. This would be expected since these parameters describe similar phenomena. There also seems to be little correlation between the response of the subject and to the absolute values of the

traces at any point. Additionally, since breath hold times between patients varied due to the subject's capability to hold her breath, features such as time to peak was difficult to extract. Therefore, in this current study we focus on the overall shape of the breath hold by looking at IE and PIE, as well as, the rise and fall time constants.

For this current analysis of dynamic features for NACT tumor response, we observed many different features and report here our most promising parameters.

6.3.2 Physiology Basis of Observed Phenomena

Like many other chemotherapeutic drugs, taxanes disrupts the cell cycle to prevent proliferation and causes apoptosis. In addition it has been shown that taxanes disrupts microvessels and decreases microvascular density.[88] However, instead of disrupting blood flow, studies in mice have suggested that the diameter of the tumor vessel and blood flow velocity inside the vessels actually increase when a tumor responds to taxanes. To explain this apparent paradox, it has been argued that solid tumors have a naturally high interstitial fluid pressure (IFP) due to the compression of blood vessels from surrounding tumor cells, the disorganized tumor vascular network, and the absence of functional lymphatics.[11, 93] Chemotherapy leads to a reduction of tumor cell density, which in turn lowers IFP, and consequently increasing blood vessel diameter and blood flow velocity. These arguments are supported by another study that demonstrated that IFP decreased after breast cancer patients who received the taxane paclitaxel.[92] Additionally, it was demonstrated that the use of taxanes can help with drug delivery to the rest of the tumor by increasing vascular surface area and essentially improving tumor response to treatment.[93] A Doppler sonography study in which 17 out of 30 patients received

paclitaxel based NACT showed an initial increase in vascularity in subjects that responded well to treatment followed by a decrease in vascularity.[97]

The results obtained in our study seem to support these findings. For example, we observed an absolute increase of m_{rise}, and m_{fall} (Table 6.3, row 3&4, column 3) which may be better explained by the cell death that occurs early during treatment, rather than anti-angiogenics effects, which would suggest a decrease in blood flow and vascular reactivity.

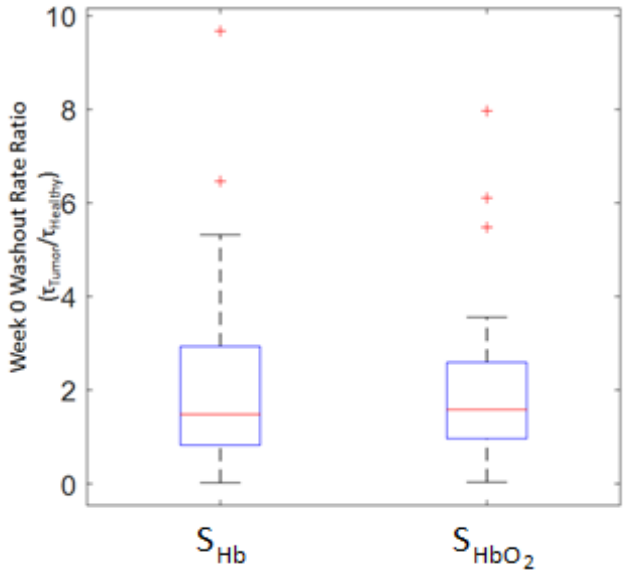


Figure 6.6 The week 0 washout rate ratio ($\tau_{Tumor}/\tau_{Healthy}$) for both S_{Hb} and S_{HbO2} were statistically significant different from the value 1 with $p < 0.01$ for each group.

The tumor region in our dynamic imaging at baseline had **a greater washout rate compared to that of the contralateral breast (see Figure 6.6)**. This finding is in agreement with other studies, based on color Doppler ultrasound, PET-MRI, and optical imaging, in which it was observed that blood flow is greater in tumor regions then in corresponding healthy breast tissue.[98-101] For example, Choe et. al. used diffuse correlation spectroscopy to observe blood flow in breast cancer patients without any

treatment. They observed that blood flow was about twice as fast in tumor regions compared to healthy tissue and compared to the contralateral breast.[100] Also, we saw increases in m_{rise}, m_{fall}, and m_{fall} for pCR subjects between the baseline and week 2 imaging time point suggesting that there may be increase blood flow in the tumor. Therefore, there is evidence in this current study to suggest that the increase apoptosis and increase IFP in responding tumors to taxane may cause the increase in m_{rise}, m_{fall}, and m_{fall} by the second week of NACT and could be physiologically relevant parameters.

6.4 Conclusions

We have developed an imaging approach to noninvasively extract information about the hemodynamic effects that occur in and around tumors of newly diagnosed breast cancer patients. In this study we demonstrate that certain features of the hemodynamic responses to a simple breath hold can be used to predict treatment outcome in patients undergoing NACT. Using dynamic optics tomographic imaging data from 34 patients, we evaluate time-traces of deoxy-hemoglobin concentrations, S_{Hb}, and oxy-hemoglobin concentration S_{HbO_2} inside the tumor. Recordings were obtained at baseline, just before treatment initiation (week 0), and 2 weeks after treatment started. Our study showed that tumors in patients with pCR (RCB-0) have statistically significant higher washout rates, m_{fall}, at week 2 compared to non-pCR (RCB-I,-II, III). Using τ_{fall} for S_{Hb} as parameter to predict patients with a pCR, we achieved a sensitivity of 92.3% and a specificity of 76%. Additionally, the difference in rise slope, m_{rise}, and fall slope, m_{fall}, between the two imaging points showed significant difference between pCR and non-pCR subjects, where pCR subjects showed an increase in each parameter for both S_{Hb}

and S_{HbO_2} traces. Finally, we determined that IE and τ_{rise} of the S_{Hb} traces are the two best features for comparing RCB-0/I/II vs. RCB III patients, i.e. the patients at highest risk for breast cancer. We were able to get 100% specificity for τ_{rise} of the S_{Hb} traces when looking at RCB-0, I, &–II versus RCB-III, which indicates that we can identify all non-responders. These are subjects that are least likely to respond NACT and hence may be prime candidates for a change of therapy.

Chapter 7: Combined Analysis

7.1 Introduction

There have been several optical and non-optical imaging studies that have observed changes of breast tumors during or after NACT. Since few patients have a pCR after treatment, the purpose of the studies was to identify the subjects who do not have a good response to treatment early during NACT. Then other options for treatment could be considered instead of continuing with a treatment that yields no result. Therefore, more personalized treatments could be developed to optimize the patient outcome of those who do not respond. [3, 48]

In Chapter 1, I go into more detail about monitoring using traditional methods. Briefly, in a meta-analysis using CE, MG, US, and MRI, that had an accuracy of 57%, 74%, 79%, 84%, respectively.[53] Additionally, some of these methods are not feasible for monitor, such as MG or MRI. MG would be dangerous for multiple imaging session due to radiation. MRI would not be feasible since it would be too expensive for multiple imaging sessions. Since traditional method lack the accuracy or the feasibility for tumor monitoring or predicting response, DOT has emerged as a novel imaging modality for tumor monitoring. DOT uses optical transmission measurements with non-ionizing, low-intensity near-infrared light to generate three-dimensional maps of various tissue parameters.

There are several advantages to using optical imaging (OI) to monitor breast tumor response over conventional imaging techniques. For example, OI uses non-ionizing light sources that allow for multiple imaging session within a short period of time. Therefore, measurements can be taken a few weeks to a few days apart from each other with no

adverse effect. OI is also sensitive to parameters that change over the course of treatment, such as tissue vascularity or hypoxia.[3, 48]

Several studies have shown that there is a change in optical properties over time during the course of neoadjuvant chemotherapy. [48, 57-60] Ueda et. al. used discriminant analysis to predict tumor response at the baseline with a sensitivity of 100% and specificity of 85.7% when combining the estrogen receptor status with baseline SO_2%. Otherwise, with just using SO_2% at baseline imaging there was a sensitivity of 75% and a specificity of 73.3%. The subjects of the study received four different treatments, some concurrent with bevacizumab and trastuzumab.[60]

Zhu et. al. found that by adding optical imaging parameters of [HbO_2]and [HbT] to their logistic regression model, there was a significant improvement to sensitivity and specificity for the pre-treatment time point. Using pathological variables alone they saw an average sensitivity of 56.8% and specificity of 88.9%. By adding [HbT] the sensitivity and specificity were 79%, and 94%, respectively. Also, by adding [HbO_2] to the model, the sensitivity became 77% and specificity was 85%. Using just the training set of data, they were able to increase the AUC by adding the [HbT] for each cycle of chemotherapy to the model, so that by just using the pathological variables the area under the curve was 87.9%, but combining the pathological variables with baseline [HbT], and cycle 1-3 measurements the AUC was 100%.[78] These analyses are based on data collected from 32 patients that received four different treatments with three patients receiving bevacizumab.[58]

Falou et. al. observed 15 patients that received several different treatments including chemoradiotherapy and was able to determine a 100% sensitivity and 83%

specificity with [Hb] and 75% sensitivity and 100% specificity with water percentage one week after the beginning of NACT. They employed univariate methods without cross validation.[86]

In the previous chapters, we looked at 3D images (static data) and 3D movies (dynamic data) individually to determine if the response of a subject could be predicted using these DOT techniques. Both studies track the tumor progression in 34 breast cancer patients receiving the same NACT regimen. The static DOT study observed how concentrations of oxy-hemoglobin, deoxy-hemoglobin, water fraction, total hemoglobin, and oxygen saturation changed over the entire course of treatment at six imaging time points and how normalized oxy-hemoglobin correlated significantly with the residual cancer burden throughout NACT. Our dynamic DOT study selected features from the hemodynamic trend that occurs as a subject held their breath and that could be used to determine the pathological response two weeks after the beginning of NACT. However, an in depth discriminant analysis of the features acquired had not been done, which is the focus of this chapter.

For the current study we employ discriminant analysis techniques and use baseline biomarkers, static DOT imaging features, and dynamic DOT imaging features to improve the accuracy, sensitivity, and specificity of our tumor prediction models. We have looked at models that separate the pCR subjects from the non-pCR subjects. Ideally, the poorly responding subjects (RCB-III) are a population of great concern, but the current number of subjects that fall into this group is 6. Therefore, discriminant analysis with this small number would be difficult to interpret. Also, since the goal of the study is to determine early on within treatment whether or not NACT is effective, we have concentrated our

feature selection to the first two imaging time points: baseline and week 2 (third cycle of taxane). Here we show that the combination of static and dynamic features yields the highest accuracy and out performs baseline biomarker analyses.

7.2 Methods

7.2.1 Patient population

There were 40 women with breast cancer imaged with four exclusions. Here, we only look at the first two time points (baseline, week 2) since the largest advantage for the patient would to be able to determine early within treatment the tumor response. A more detailed description of the subject population can be found in Table 2.3.

7.2.1 Pathological, Static, and Dynamic Parameters

There were five data sets we incorporated for the current study: pathology, baseline static data, week 2 static data, baseline dynamic data, and week 2 dynamic data. These data were obtained from the results of the previous chapters. For the pathological biomarkers, the estrogen receptor status (ER), progestogen status (PR), and Ki-67 protein percentage (Ki67) were acquired before the initiation of NACT. Four patients were missing their baseline Ki67 tests due to outside diagnosis and testing. The static DOT parameters include concentration of [HbO2], [Hb], WF, [HbT], SO2%, and the normalized concentrations by both the healthy breast (T/H) and the non-tumor region in the tumor bearing breast (T/nT). Only week 2 features have the percent change from baseline imaging. Additionally, the percent change from the baseline was found for the normalized values as well. These parameters are described in more detail in chapter 5. The dynamic parameters are taken from the hemodynamic time traces of the percent change in [HbO_2]

(S_{HbO2}) and [Hb] (S_{Hb}) during a breath hold. Some of the features include the rise slope, fall slope, rise rate, washout rate and maximum peak value. The calculations and descriptions of the dynamic features can be seen in chapter 4. A list of all parameters are shown in Table 7.1.

Table 7.1 Summary of features where X could be any of the features above and Y is any feature above or any feature involving X.

Pathological features	
ER	Estrogen Receptor percentage
PR	Progestogen Receptor percentage
Ki67	Ki-67 protein percentage
Static Features	
[HbO_2]	Oxy-hemoglobin Concentration
[Hb]	Deoxy-hemoglobin Concentration
WF	Water Fraction
[HbT]	Total hemoglobin Concentration
SO_2%	Oxygen Saturation
$X_{(T/H)}$	Feature X normalized to Healthy Breast
$X_{(T/nT)}$	Feature X normalized to non-Tumor region in Tumor bearing breast
%ΔY	Percent change of feature Y from baseline imaging

Dynamic Features	
m_{Rise}	Rise Slope
m_{Fall}	Fall slope
τ_{Rise}	Rise rate
τ_{Fall}	Fall rate
IE	Initial Enhancement
PIE	Post Initial Enhancement
S_{Peak}	Peak value
NMPV	Normalized maximum peak value
CC	Correlation Coefficient
X for S_{Hb}	Parameter X for deoxyhemoglobin dynamic curves
X for S_{HbO2}	Parameter X for oxyhemoglobin dynamic curves
ΔY	Change in Y from baseline imaging
%ΔY	Percent change of Y from baseline imaging

7.2.2 Discriminant Analysis

Discriminant analysis was performed using MatLab. Both univariate and multivariate methods were employed for calculating the accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for determining whether the patient had a pCR or non-pCR to treatment. Sensitivity in this study represents the probability that a positive predicted condition is a true pCR. Conversely, specificity represents the probability that a negative predicted condition is a true non-pCR.

The data was divided into testing and training data where 25% of the data was used for testing. The univariate and multivariate discriminant analysis was repeated 20 times with a different random portion of the data removed for testing. Then the mean and standard deviation were calculated for the accuracy, sensitivity, specificity, PPV, and NPV.

For the univariate method, the cutoff point/threshold for was taken at the Youden index (YI) of the training data and then applied to the testing data. As for the multivariate method both linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) were employed. These methods were used for the baseline and week 2 time points, and the static and dynamic data. Last, the static and dynamic features were both used for LDA and QDA.

The feature for the univariate analysis with the largest YI are was reported. For the multivariate method, LDA and QDA were performed for every combination of features between two sets of data and the combination with the highest YI was reported.

7.3 Results

7.3.1 Univariate Analysis

First, we looked at the ROC curves of the just the pathological baseline biomarkers to determine the standard ROC that the physicians would have for this particular study. Figure 7.1 shows the ROC curves for ER, PR, and Ki67 that have an area under the curve (AUC) of 0.89, 0.85, and 0.53, respectively.

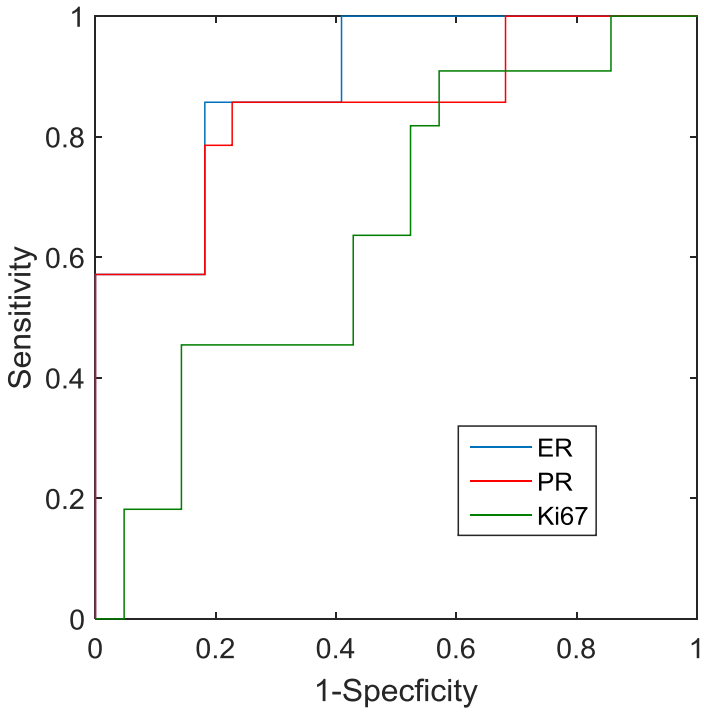


Figure 7.1 ROC curves of pathologic biomarkers ER, PR, and Ki67.

ER has the highest AUC with a sensitivity of 85.7%, a specificity of 81.8%, and an accuracy of 83.3%. The sensitivity analysis at the YI for the biomarkers can be seen in Table 7.2. Additionally, these pathological biomarkers underwent univariate analysis with cross validation (Table 7.3).

Table 7.2 Sensitivity analysis of ROC curves in Figure 7.1 for ER, PR, and Ki67.

Biomarker	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
ER	0.89	85.71%	81.82%	83.3%	75.0%	90.0%
PR	0.85	85.71%	77.27%	80.6%	70.6%	89.5%
Ki67 (n=32)	0.53	90.91%	42.86%	59.4%	45.5%	90.0%

First, univariate analysis was performed to see if a single variable would be viable enough for tumor response prediction at the first two imaging time points during NACT. Table 7.3 shows the features for each set of data that had the highest Youden index. Out of the data sets shown, pathology had the highest YI by using the ER percentage.

Consequently, ER had the highest accuracy (72.8%) and sensitivity (64.2%). However, the sensitivity and accuracy for any of these parameters is <75%. Therefore, multivariate analysis was performed to see if the accuracy, sensitivity, and specificity could be increase.

Table 7.3 Summary of univariate analysis for each data set with the feature with the greatest Youden index.

Data Set	*Feature*	*Accuracy*	*Sensitivity*	*Specificity*	*PPV*	*NPV*
Pathology	ER	72.8%±11.7%	64.2%±14.6%	89.6%±14.8%	91.4%±12.5%	56.9%±17.4%
Static Baseline	[Hb]	46.1%±15.4%	25.0%±44.4%	75.0%±44.4%	51.1%±14.9%	44.4%±15.7%
Static Week 2	$\%\Delta[HbO_2]_{(T/nT)}$	71.3%±11.9%	53.3%±20.7%	96.6%±8.2%	95.3%±12.6%	61.2%±14.7%
Dynamic Baseline	τ_{Rise} for S_{HbO2}	50.0%±13.7%	21.8%±12.8%	96.5%±9.0%	88.9%±27.4%	43.2%±15.2%
Dynamic Week 2	PIE for S_{Hb}	70.0%±14.5%	62.3%±22.1%	85.5%±14.2%	87.0%±12.5%	61.2%±22.2%

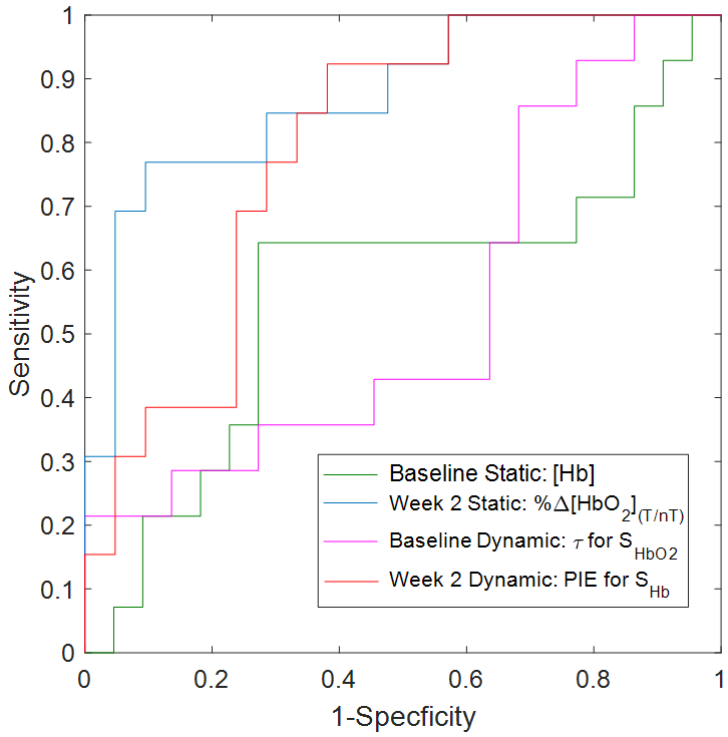


Figure 7.2 ROC curves of best DOT features from univariate analysis.

Figure 7.2 shows ROC curves for the best DOT features from the univariate analysis. Here the highest area under the curve is the week 2 static feature of

$\%\Delta[HbO_2]_{(T/nT)}$. This features has a sensitivity of 76.9% and specificity of 90.5% (see Table 7.4). Additionally, $\%\Delta[HbO_2]_{(T/nT)}$ has the highest accuracy of the features (85.5%), which is higher than the ER ROC analysis. However, $\%\Delta[HbO_2]_{(T/nT)}$ has a lower sensitivity than ER.

Table 7.4 Summary of sensitivity analysis of the best select features from the univariate analysis seen in Table 7.3

Feature	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
Baseline Static:[Hb]	0.56	64.3%	72.7%	69.4%	60.0%	76.2%
Week 2 Static: $\%\Delta[HbO_2]_{(T/nT)}$	0.87	76.9%	90.5%	85.5%	83.3%	86.4%
Baseline Dynamic: τ_{Rise} for S_{HbO2}	0.54	21.4%	100.0%	69.4%	100.0%	66.7%
Week 2 Dynamic: PIE for S_{Hb}	0.79	92.3%	61.9%	73.5%	60.0%	92.9%

7.3.2 Multivariate Analysis

Table 7.5 Summary of LDA and QDA for each data set with the features with the greatest Youden index.

Data Set	*Type*	*Feature I*	*Feature II*	*Accuracy*	*Sensitivity*	*Specificity*	*PPV*	*NPV*
Pathology	LDA	ER	PR	72.8%±15.5%	80.0%±20.3%	68.7%±18.2%	62.7%±23.2%	83.7%±14.6%
	QDA	ER	PR	75.0%±21.3%	78.3%±37.5%	57.5%±43.8%	72.5%±25.0%	82.3%±35.2%
Static Baseline	LDA	$[HbT]_{(T/nT)}$	$[HbO_2]_{(T/H)}$	70.6%±14.1%	43.4%±30.5%	89.1%±19.5%	79.6%±32.1%	72.6%±14.5%
	QDA	$[Hb]_{(T/nT)}$	$[HbT]_{(T/H)}$	67.8%±17.6%	42.8%±21.2%	87.7%±18.8%	77.6%±33.1%	64.4%±20.6%
Static Week 2	LDA	$SO_2\%_{(T/nT)}$	$\%\Delta[HbO_2]_{(T/H)}$	82.5%±16.4%	80.8%±28.2%	85.5%±18.4%	73.7%±33.2%	90.3%±14.5%
	QDA	$\%\Delta SO_2\%$	$\%\Delta WF_{(T/nT)}$	81.3%±11.1%	84.4%±22.5%	83.7%±15.5%	72.3%±24.7%	90.0%±13.7%
Dynamic Baseline	LDA	Fs S_{Hb}	NMPV for S_{Hb}	65.6%±16.5%	48.4%±27.0%	83.6%±19.7%	72.1%±27.8%	68.1%±19.3%
	QDA	Fs for S_{Hb}	τ_{Rise} for S_{Hb}	78.3%±14.2%	92.7%±19.7%	70.2%±23.3%	73.1%±21.5%	93.3%±16.7%
Dynamic Week 2	LDA	ΔFs S_{Hb}	PIE for S_{Hb}	88.8%±13.4%	81.3%±28.5%	92.1%±12.4%	82.1%±29.4%	91.7%±11.5%
	QDA	τ_{Fall} for S_{Hb}	ΔS_{Peak} for S_{HbO2}	88.1%±7.6%	89.0%±16.3%	87.9%±10.4%	82.6%±15.9%	93.0%±10.0%

In Table 7.5 is a summary the features that had the highest YI of the discriminant analysis for each data set using both LDA and QDA. The QDA for dynamic week 2 data has the highest YI, where accuracy is 88.1%, sensitivity is 89.0%, specificity is 87.9%, PPV is 82.6%, and NPV is 93.0%. However, the LDA for dynamic week 2 had a higher

accuracy (88.8%). Both the static and dynamic week 2 data set had the highest accuracies (>80%) among the analyzed data sets.

The best performing test for the baseline data sets was the dynamic data set with QDA where the accuracy was 78.3%, the sensitivity was 92.7% (the largest sensitivity from Table 7.5), and a specificity of 70.2% .

This is a large increase in accuracy compared to the univariate analysis. LDA and QDA yield similar results for each of the data sets, in which QDA only occasionally increases accuracy or the YI.

7.3.3 Multivariate Analysis with Pathology

Table 7.6 Summary of LDA and QDA for pathology combined with DOT features.

Data Set	*Type*	*Pathology Feature*	*DOT Feature*	*Accuracy*	*Sensitivity*	*Specificity*	*PPV*	*NPV*
Static Baseline	LDA	ER	SO_2%	75.6%±15.1%	81.9%±24.8%	75.8%±15.0%	68.2%±19.5%	84.7%±18.7%
	QDA	ER	[HbT]	75.6%±17.1%	82.0%±21.1%	75.9%±23.8%	68.0%±26.6%	86.8%±16.4%
Static Week 2	LDA	PR	%Δ[HbO2]	85.0%±11.2%	86.8%±25.2%	83.9%±14.0%	75.5%±19.3%	92.9%±12.1%
	QDA	ER	$\%\Delta[HbT]_{(T/H)}$	85.0%±10.4%	80.3%±18.0%	88.2%±15.3%	82.9%±21.0%	88.2%±10.2%
Dynamic Baseline	LDA	PR	m_{Fall} for Hb	82.2%±12.2%	86.3%±18.8%	82.0%±18.7%	76.7%±22.7%	90.1%±14.4%
	QDA	ER	IE for Hb	78.9%±12.4%	85.6%±17.4%	76.3%±18.9%	73.7%±21.2%	86.5%±15.7%
Dynamic Week 2	LDA	ER	Δm_{Fall} for Hb	85.6%±11.7%	81.5%±22.2%	90.5%±12.6%	85.8%±18.8%	87.5%±16.7%
	QDA	PR	$\%\Delta\tau_{Fall}$ for Hb	84.4%±9.0%	76.3%±24.1%	91.6%±11.1%	86.7%±17.4%	86.0%±12.4%

Since there was still room for improvement for the YI and accuracy, we combined each of the DOT parameters with the pathology of the subject. Table 7.6 shows the combination of features with the highest YI for each data set. The best performing test was the LDA of pathology combined with dynamic week 2. This combination has an accuracy of 85.6% with a sensitivity and specificity of 81.5% and 90.5%, respectively. The next best performing test is the QDA for the combination of pathology features and static week 2 features, where the accuracy, sensitivity, and specificity are 85.0%, 80.3%, and 88.2%, correspondingly. The best performing pathology and baseline data set

combination was LDA for dynamic week 2 with an accuracy of 82.2%, sensitivity of 86.3%, and specificity of 82.0%.

The top results from the multivariate analysis are an improvement from the univariate analysis of the pathology.

7.3.4 Combining Baseline and Week 2 features

Table 7.7 Summary of LDA and QDA for baseline versus week 2 features.

	Type	*Baseline Feature*	*Week 2 Feature*	*Accuracy*	*Sensitivity*	*Specificity*	*PPV*	*NPV*
Static	LDA	$WF_{(T/H)}$	$\%\Delta[HbT]_{(T/nT)}$	81.3%±13.1%	81.7%±23.5%	84.2%±13.4%	73.8%±22.6%	87.7%±15.4%
	QDA	$[Hb]_{(T/H)}$	[HbT]	75.0%±18.6%	82.0%±31.6%	75.5%±20.4%	65.1%±28.5%	87.5%±20.8%
Dynamic	LDA	NMPV for S_{HbO2}	Δm_{Fall} for S_{Hb}	86.3%±10.7%	76.3%±32.0%	90.7%±11.1%	77.2%±29.9%	90.4%±11.4%
	QDA	m_{Fall} for S_{Hb}	τ_{Fall} for S_{Hb}	83.1%±10.2%	81.7%±22.1%	85.5%±11.3%	71.9%±22.1%	89.0%±14.4%

In order to see if the static or dynamic data sets could be used individually, LDA and QDA were used to combine baseline a week 2 features. The highest YI index was the dynamic QDA test where the accuracy was 83.1%, sensitivity was 81.7%, and specificity was 85.5% using the Fs for S_{Hb} and τ_{Fall} for S_{Hb} as the features. However, the highest accuracy and specificity was from the dynamic LDA test (accuracy 86.3%, specificity 90.7%).

Besides the baseline parameters, this analysis does not yield higher accuracies or YI from the multivariate analysis using only one data set at a time.

7.3.5 Combining Static and Dynamic features

Table 7.8 Summary of LDA and QDA for static versus dynamic features.

	Type	*Static*	*Dynamic*	*Accuracy*	*Sensitivity*	*Specificity*	*PPV*	*NPV*
Baseline	LDA	$[HbO_2]_{(T/H)}$	NMPV for S_{HbO2}	63.3%±17.3%	39.9%±34.5%	83.5%±17.7%	55.2%±35.3%	71.2%±20.5%
	QDA	$SO_2\%_{(T/nT)}$	Fs for S_{Hb}	67.2%±13.2%	54.0%±27.7%	82.1%±21.4%	73.3%±24.3%	71.5%±19.5%
Week 2	LDA	$\%\Delta[HbO_2]_{(T/H)}$	τ_{Fall} for S_{Hb}	91.3%±10.8%	89.7%±15.6%	95.0%±11.3%	91.7%±19.3%	92.0%±11.6%
	QDA	%Δ[HbO2]	τ_{Fall} for S_{Hb}	93.1%±8.6%	87.1%±25.4%	94.4%±10.5%	90.6%±17.4%	94.8%±9.9%

The static and dynamic parameters were combined and the test with the highest YI was LDA for week 2, but the highest accuracy was QDA for week 2. The LDA for week 2 had an accuracy of 91.3%, a sensitivity of 89.7%, and a specificity of 95.0%. Also, the week 2 LDA test had a PPV of 91.7% and a NPV of 92.0%. As for the QDA, there was an accuracy of 93.1%, which is the highest accuracy out of all the previous discriminant analyses. The combination of data sets can be seen in Figure 7.3, where a non-validated QDA was performed. Here the accuracy, sensitivity, specificity, PPV, and NPV were 94.1%, 92.3%, 95.2%, 92.3%, and 95.2%, respectively.

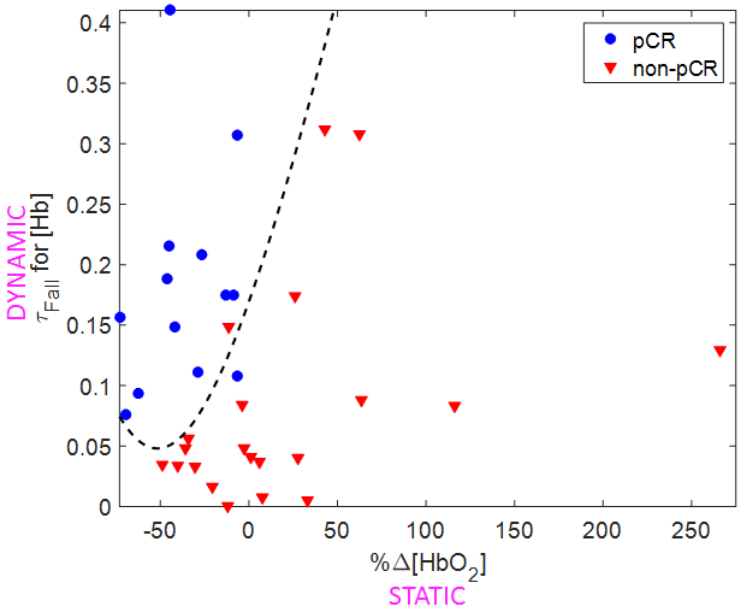


Figure 7.3 QDA without validation of the week 2 features from the static and dynamic data sets. The accuracy, sensitivity, specificity, PPV, and NPV were 94.1%, 92.3%, 95.2%, 92.3%, and 95.2%, respectively.

7.4. Discussion

The univariate analysis showed that the pathology had the best YI, but was not clinically useful. Yet, the analysis gave us a standard in which we needed to exceed in

which ER status had a sensitivity of 64.2% and a specificity of 89.6%. The multivariate analysis for both week 2 static and dynamic data sets performed much better than the univariate analysis. The week 2 data sets had accuracies that were all over >80%. From the baseline data sets the dynamic QDA had the highest YI and sensitivity (92.7%), which may be helpful at determining pCR at the baseline imaging. However, the multivariate analysis was most accurate when the week 2 data was taken into consideration.

The best baseline testing was the combination of pathology and dynamic features, which yielded an accuracy of 82.2% for LDA (see Table 7.6). Here, since the accuracy, sensitivity, specificity, PPV, and NPV are all >75.0%, this test may be viable for baseline screening of NACT effectiveness before the first dose is administered. However, the analysis clearly shows that the week 2 time point yields better testing options that give higher accuracies, sensitivities, etc. The best week 2 performing combination would be between the static and dynamic features, where an accuracy of 91.3% was acquired from LDA, and an accuracy of 93.1% was acquired from QDA. Therefore, the QDA model may be used tumor response prediction (see Figure 7.3). Future studies could consist of using this model to determine which patients will not have a pCR and possibly add or change the agents used from their current treatment protocol.

The physiological features that yielded the highest accuracy from our discriminant analysis test were $\%\Delta[HbO_2]$ and the washout rate τ_{Fall} at the week 2 time point. The physiological basis of each of these parameters has been explained before in sections 5.4 and 6.4 of this thesis. Briefly, there is a tendency for pCR subjects to have a decrease in $[HbO_2]$ levels by the second week of NACT initiation. Also, we have observed that pCR subjects have a higher washout rate by week 2 compared to non-pCR subjects. As a

result, the pCR subjects tend to cluster in the top left portion of the graph that compares these two features (see Figure 2). While the non-pCR subjects tend to have little change or an increase in [HbO_2] levels, as well as, a slower washout rate at the week 2 time point. Therefore, the combination of the features yields a very good prediction model for tumor response to NACT.

The result with the highest YI was the LDA with the week 2 static and dynamic data combined. With an accuracy of 91.3%, a sensitivity of 89.7%, and a specificity of 95.0%, our model outperformed the traditional methods seen in Table 1.6. Even MRI, which had the accuracy of 84% was still lower than the current results we have obtained. Therefore, using dynamic and static information can plausibly predict tumor response to NACT treatment.

Overall, the discriminant analysis for pathology, static features, and dynamic features have yielded promising results that could potentially be used in interventional studies to improve the outcome of the patient, based on their DOT imaging. Especially, combining the static and dynamic capabilities of our imaging system to obtain different features to use in simple discriminant analysis techniques has acquired a prediction accuracy as high as 93.1%.

Chapter 8

8.1 Summary

Our current DOT breast imager is capable of dual breast dynamic and static imaging. Therefore, there is a vast array of information we can utilize to observe tumor vascular changes. The static images are capable of locating the tumor position, as well as showing changes over treatment that correspond to the pathological response of the patient. Additionally there are visible changes in the dynamic traces of the patient breath hold that also can be analyzed to determine if the subject has a good response. The combination of this information has been utilized to further assist in determining response. Ultimately the goal was to use the week 2 information to predict whether early in the treatment if there is a response and change the treatment plan accordingly to optimize the patients' outcomes.

There were 40 subjects have been recruited for this current study with 36 subjects that have been used for the overall analysis of this NACT monitoring study. The number of patients in our study that have a pCR was about 38.9%, which corresponds to similar numbers in the literature.[1, 3] Our analysis has shown promising results, with significant difference in percent change of [HbT] at the week 2 time point. Our results correspond with previous studies where pCR patients show a drop in [HbT] after the first few weeks of treatment. However, in other studies, subjects received different chemotherapy regimens or have taken additional cancer drugs.[10-12, 15-20] Here, we look at subjects that have all undergone the same treatment, throughout the length of their therapy.

The static images with different chromophore concentrations that were observed indicate that the baseline and week 2 time points were the most opportune chance for

determining a pCR or not to treatment. We also observed that baseline $[HbO_2]_{(T/H)}$ significantly correlates with RCB score. Additionally, several parameters for week 2 showed significant correlations. When the percent change from the baseline was calculated, $\%\Delta[HbO_2]_{(T/nT)}$ show significant correlation throughout each of the imaging time points with the RCB score.

The hemodynamic effects visualized in this study were acquired by using non-invasive methods to understand the vasculature of healthy and diseased tissue. Our analysis shows how the hemodynamics of breast cancer changes as a patient receives treatment over two weeks. This information could be beneficial to physicians as a method of tumor monitoring and treatment evaluation. For pCR subjects, we observed that there is an increase in the washout rate, fall slope, and rise rate by the second imaging time point. The possible physiological reason could be the increase in apoptosis which decreases IFP and allows an increase of blood flow.

Last by combining the static and dynamic data sets and performing simple discriminant analysis methods, we were able to obtain an accuracy as high as 93.1% at determine pCR subjects from non-pCR subjects.

8.2 Future Work

8.2.1 Hardware Development

There are several possible modifications to the current imaging system that could be made to increase the accuracy and quality of the images and the overall amount of information acquired. There could be the addition of wavelengths that could assist with image reconstructions to better visualize chromophore concentrations, such as lipid.

Also, there could be a frequency domain (FD) component added to increase the accuracy of the scattering coefficient and the overall reconstruction image. Last, the patient interface could undergo minor modification to help to ensure contact.

Currently, lipid cannot be reconstructed using the current wavelengths that the system has. While, important chromophores such as oxy and deoxy-hemoglobin can still be visualized, adding another feature to the analysis could help determine the response of the subject over the course of treatment. Additionally, having more laser wavelengths than chromophores could increase the accuracy of the reconstruction algorithm by decreasing the ill-posedness of the solution.[102]

Since the breast imager is a continuous wave system, there is no phase component recorded. The phase component of the signal usually helps distinguish between the absorption and scattering coefficient[102]. A FD component could be added in order to get a baseline image of the subject. Then the system could be switched to CW mode and continue with the dynamic imaging.[32, 35, 36] Then the baseline FD measurement can be used to determine the initial scattering coefficient and could be used for the static measurements.

For the patient interface, there could be pressure sensors setup along the rings of the interface. These pressure sensors could be coupled with the computer and give a binary signal in which it indicates good or bad contact with the detector fibers. Occasionally, proper contact with the patient interface is hard to determine. The pressure sensors could take out some of the guess work for optimal patient positioning and increase the quality of the image, since (as shown in chapter 3) poor contact leads to artifacts in the image.

8.2.2. Static Imaging Improvements

The current static reference poses some logistical problems. The liquid phantoms only last about a month before they are no longer useful. The intralipid for these phantoms are also relatively expensive and could cost about $80 per phantom. Solid phantoms would be more cost effective. The upfront cost (~$3,000) would be relatively expensive, but would pay for its self over time, since it would not need to be replaced. The solid phantom could be shaped to each of the rings of the patient interface, which would prevent any contact issues with the detectors. Also, there would be no variability between phantoms since the same one would be used for all imaging. Additionally, the solid phantoms would be key in multi-center studies in order to remove the variability of liquid phantom preparation by different groups. Therefore, with solid phantoms, the accuracy of the static images could increase and possible artifacts could be avoided.

8.2.3 Dynamic Imaging

One issue that the dynamic portion of the study has run into is knowing when a patient holding their breath. We depend on the honesty of the patient to tell us when they begin breathing again, but sometimes it is apparent that the patient is not holding her breath. Also, when they begin or end the breath hold is up to them and we cannot always observe these timings. To fix this issue, there may be two possibilities: have the patient breath in different gasses or have the patient wear a strain gage that tracks their breathing.

There has been a study where the patient breathes in a different combination of gasses and the tumor can be distinguished.[103] This would be one way of controlling the experiment and making sure that each patient underwent the same protocol. With the

breath hold method, each patient holds their breath for a different length of time, but there would be more control over the timing with the inhalation of timed gasses. However, this does introduce and extrinsic element that patients may not appreciate.

Also, there is a simple strain gage device that can be placed around the chest (below the breasts) that can measure the expansion and relaxation of the muscles around the lungs. The strain gage device can provide a signal to the computer and recorded in tandem with the optical imaging. Hence, the breathing of the patient and the optical imaging data could be digitally match up and possibly automatically segmented to show just the breath hold data. This is a simple solution that would not add an extrinsic factor to the patient, such as the different gasses. However, the downside would be that the breath hold times could still range from subject to subject.

8.2.4 Patient comfort

Patient comfort is often neglected in scientific articles, but I feel that it is somewhat worth mentioning since the system was designed to allow the patient to be as comfortable as possible. Our current system and study can provide a comfortable and fairly quick means of imaging patients with breast cancer. There is no need for painful compression that is required by mammography. Also, the actual imaging takes about 10-12 minutes (maybe 20 minutes with setup time), which would be much quicker than a patient would experience in a MRI system. Additionally, when we examine the hemodynamic phenomena caused by the tumor vasculature, we look at intrinsic contrasts that occur during a simple breath hold so there is no need for injections of contrast agents.

I believe these factors have had a large impact on our capability to maintain constant accrual of patients on the study and to have them return for all or most their

imaging sessions. However, there are still improvements that can be made to ensure the comfort of the patient. First, one of the biggest complaints I get about the system is that the patient's back or neck will hurt from the hunch over position that they are in. There was even a point in the study where we stopped using a chair and had the patients stand throughout the imaging since they were uncomfortable and there was a lack of space in the new imaging room. However, this problem is still occurring, especially when the patient already had a preexisting condition that involves shoulder or back pain. This problem could be eradicated with a system that was set up like a bed. A bed interface would be more comfortable, could possible lead to better contact with the patient interface, and lead to less motion from the patient who may adjust themselves during imaging to be more comfortable.

8.2.5 Imager Software Development

Currently, the DDOT breast imager has a LabView GUI that controls the system. This system is difficult to navigate. An engineer that has been trained in LabView can use the DDOT GUI interface. However, for a research coordinator, clinician, or physician the interface could be confusing and takes a while to train a single person on how the program works. Therefore, to increase the ease of use of the system the GUI should be rearranged. The best way might be to have a series of windows that prompt the user to enter the necessary information. For example, at the beginning the program, it prompts the user to enter the number of sources and detectors for the subject. Afterwards, the program could ask if the patient is in position and then automatically configure the gain bits. Then the subject information could be entered and then automatically begin imaging.

Overall, a more user friendly interface is needed in order to increase personnel that can operate the system. For multi-center studies this would be essential.

8.2.5.1 Pre-processing and Post-processing

The short comings of some of the pre- and post-processing techniques have been explained in previous chapters. Yet, I would like to mention that there are still work to be done with this regard. Overall, these pre- and post-processing tools could be placed into a single MatLab program or GUI for both the dynamic and static data, which would increase the ease of use and decrease the time spent on processing.

8.2.5.2 Real-Time Imaging

A key facet and long term goal of DOT imaging would to obtain real time images. The current set up lets us take the data a reconstruct the images over a day. However, with most imaging systems such as US, imaging can be viewed in real time. The limiting factor for DOT are the reconstruction algorithms. Our current study reconstructs 3D images and custom designed meshes. However, fast 2D images may be possible to procure. Since the rings of the breast imager are fixed, the mesh for each ring will always remain the same. Hence, 2D meshes could be generated once and used for real-time reconstruction. Also, 2D reconstructions are much faster than 3D reconstructions since there are many less units of space needed to solve the inverse problem. Additionally, to increase the speed the first three or four iterations could be used as the final image (instead of waiting for 9-10 iterations). As result, we could get images within minutes instead of hours for a single imaging. Later, full tomographic images could be created in order to increase the accuracy and understanding of the data (tumor location, hemodynamic effects, etc.). The real-time imaging could aid in quick diagnosis and tumor

response predictions that could help the physician and patient move forward with treatment.

8.2.6 Interventional Studies

By using the current data set and developing a model for discriminating the response groups at the baseline or early treatment time point such as the second week, interventional studies could be developed. The interventional study could consist of baseline imaging and initial assessment. We have develop discriminant analysis models that can have a sensitivity as high as 94.1% at determining a pCR. Therefore, these subjects would unlikely need further monitoring.

For the subjects that seem like they will not have a pCR, DOT can further monitor at week 2 to check for a pCR status. Also, we can compare the DOT values to the linear regressions of the static data and estimate the RCB score at each time point. Further analysis would be needed to determine specifically the RCB-III subjects (subjects with no response) since the aforementioned study had only 6 subjects in this group.

8.2.7 Multi-Center Clinical Trials

The NACT DOT imaging study is considered a Phase II clinical trial, where a set of subjects are studied for an initial analysis of an experimental imaging system. The next step would be to go to a Phase III clinical study where there would be multiple imaging systems and collaborators at other institutions or hospitals involved. The advantage of this study is threefold. First, this would give us a different subject population that we were not capable of reaching at New York Presbyterian Hospital, which would valuable to know if the imaging system could work on this subject population as well. Second, that the imaging system is user independent and that similar results could be obtained from

different locations, which would show the consistency of the system. Last, patients could be recruited regularly and would yield a much bigger data set than a single center study.

Organizing and performing a successful multi-center trial could yield results that could help obtain FDA approval and ultimately allow for clinicians to use the DOT system within the hospital to help monitor breast cancer.

8.2.8 Modelling Hemodynamic Trends

Now that we have developed a method to acquire and analyze the hemodynamic traces from the breath hold imaging, we can look at ways to model these trends. As I have looked at these breath hold traces, I have noticed a variety of differences among the curves. Yet, we do not understand fully what causes these different features to occur. Therefore, by modelling theses traces using mock circuit designs, such as the Windkessel model, we can possibly mimic the trends occurring in the vasculature and come to a better understanding of the physiology. Last, once we have developed these models, we can use them as another method to predict the outcome of NACT based on the resistance or capacitance of different sections of the model.

8.2.9 Additional Data Analysis

Due to the large quantity of data, there is still data that needs to be processed. For example, the dynamic data for all time points must still be reconstructed and analyzed to view any possible trends in the hemodynamics over time. Additionally, discriminant analysis can be performed on all time points and prediction models could be developed.

8.3 Conclusion

Both the optical properties and hemodynamic trends of breast tumors can indicate how a patient will respond to NACT. Additionally, DDOT can easily be employed to safely and inexpensively image a patient several times over the course of treatment and their tumors can be monitored throughout each stage. We have observed that the $\%\Delta[HbO2]_{(T/nT)}$ correlates with the RCB score of the patient population at each imaging time point. The dynamic study shows that there is an increase in rise slope, fall slope, and washout rate in pCR subjects by week 2, which could be explained by the increase in blood flow due to the apoptosis of the tumor and decrease of IFP. Additionally, combining the static and dynamic data, we have developed prediction models that are capable of predicting the response of the subject at the week 2 imaging time point with an accuracy of 94.1%. Therefore, DDOT static and dynamic capabilities can be utilized to observe malignant tumors, predict treatment outcome, and monitor tumor response throughout NACT. As a result, DDOT is a viable imaging modality for an area of breast cancer treatment that does not currently exist.

References

1. Esserman, L.J. Neoadjuvant Chemotherapy for Primary Breast Cancer: Lessons Learned and Opportunities to Optimize Therapy. *Annals of Surgical Oncology* **11**, 3S-8S (2003).
2. Schott, A.F. et al. Clinical and radiologic assessments to predict breast cancer pathologic complete response to neoadjuvant chemotherapy. *Breast Cancer Research and Treatment* **82**, 231-238 (2005).
3. Tromberg, B.J. et al. Assessing the future of diffuse optical imaging technologies for breast cancer management. *Medical Physics* **35**, 2443-2541 (2008).
4. American Cancer Society. Breast Cancer Facts & Figures 2013-2014. *Atlanta: American Cancer Society, Inc.* (2013).
5. SEER Cancer Statistics Factsheets: Breast Cancer. *National Cancer Institute. Bethesda, MD,* http://seer.cancer.gov/statfacts/html/breast.html.
6. Grimsey, E. in Breast Cancer Nursing Care and Managment, Edn. 2. (ed. V. Harmer) 1-18 (John Wiley & Sons, Ltd, 2010).
7. Johnson, M.C. in Managment of Breast Disease. (eds. I. Jatoi & M. Kaufmann) 1-32 (Springer, 2010).
8. Froyd, H.E. & Harmer, V. in Breast Cancer Nursing Care and Management. (ed. V. Harmer) 19-34 (John Wiley & Sons, Ltd., 2010).
9. Bae, Y.K. & Gabreilson, E.W. in Managment of Breast Disease. (eds. I. Jatoi & M. Kaufmann) 169-179 (Springer, 2010).
10. Bergers, G. & Benjamin, L.E. Tumorigenesis and the angiogenic switch. *Nat Rev Cancer* **3**, 401-410 (2003).
11. Minchinton, A.I. & Tannock, I.F. Drug penetration in solid tumours. *Nature Reviews* **6**, 583-592 (2006).
12. Hooley, R.J., Andrejeva, L. & Scoutt, L.M. Breast Cancer Screening and Problem Solving Using Mammography, Ultrasound, and Magnetic Resonance Imaging. *Ultrasound Quarterly* **27**, 23-46 (2011).
13. Urban, L. & Urban, C. Role of Mammography versus Magnetic Resonance Imaging for Breast Cancer Screening. *Curr Breast Cancer Rep* **4**, 216-233 (2012).
14. Esserman, L.J., Jr., I.M.T. & Reid, B. Overdiagnosis and Overtreatment in Cancer. *JAMA* **310**, 797-798 (2013).
15. Miller, A.B. et al. Twenty five year follow-up for breast cancer incdence and mortality of the Canadian National Breast Screening Study: randomised screening trial. *BMJ* **348** (2014).
16. Kolb, T.M., Lichy, J. & Newhouse, J.H. Comparison of the Performance of Screening Mammography, Physical Examination, and Breast US and Evaluation of Factors that Influence Them: An Analysis of 27,825 Patient Evaluation. *Radiology* **225**, 165-175 (2012).
17. , Vol. 2014 (American Cancer Socity, 2014).
18. Weigel, S., Biesheuvel, C., Berkemeyer, S., Kugel, H. & Heindel, W. Digital mammography screening: how many breast cancer are additionally detected by bilateral ultrasound examination during assessment? . *Eur Radiol* **23**, 684-691 (2013).
19. Pilewskie, M. & King, T.A. Magnetic Resonance Imaging in Patients with Newly Diagnosed Breast Cancer. *Cancer* (2014).
20. Vreugdenburg, T.D., Willis, C.D., Mundy, L. & Hiller, J.E. A systematic review of elastography, electrical impedance scanning, and digital infrared thermography for breast cancer screening and diagnosis. *Breast Cancer Res Treat* **137**, 665-676 (2013).
21. Calhoun, B.C. & Livasy, C.A. Mitigating Overdiagnosis and Overtreatment in Breast Cancer: What is the Role of the Pathologist. *ARch Pathol Lab Med* (2014).
22. Flexman, M.L. et al. Digital optical tomography system for dynamic breast imaging. *Journal of Biomedical Optics* **16** (2011).
23. Flexman, M.L. et al. Optical biomarkers for breast cancer derived from dynamic diffuse optical tomography. *Journal of Biomedical Optics* **18**, 096012-096012 (2013).
24. Liedtke, C. et al. Response to Neoadjuvant Therapy and Long-Term Survival in Patients With Triple-Negative Breast Cancer. *Journal of Clinical Oncology* **26**, 1275-1281 (2008).
25. Wang, J. et al. In vivo quantitative imaging of normal and cancerous breast tissue using broadband diffuse optical tomography. *Medical Physics* **37**, 3715-3724 (2010).

26. Zhu, Q. et al. Benign versus malignant breast masses: optical differentiation with US-guided optical imaging reconstruction. *Radiology* **237**, 57-66 (2005).
27. Scholl, S.M. et al. Neoadjuvant versus adjuvant chemotherapy in premenopausal patients with tumours considered too large for breast conserving surgery: preliminary results of a randomised trial: S6. *Eur J Cancer* **30A**, 645-652 (1994).
28. Wang, L.V. & Wu, H.-I. Biomedical Optics: Principles and Imaging. (John Wile & Sons, Inc., Hoboken, New Jersey; 2007).
29. Prahl, S.A., Vol. 2016 (
30. Hielscher, A.H. et al. Near-Infrared Diffuse Optical Tomography. *Disease Markers* **18** (2002).
31. Leff, D.R. et al. Diffuse optical imaging of the healthy and diseased breast: A systematic review. *Breast Cancer Res Treat* **108**, 9-22 (2008).
32. Fang, Q. et al. Combined Optical Imaging and Mammography of the Healthy Breast: Optical Contrast Derived from Breast Structure and Compression. *IEEE Transactions on Medical Imaging* **28**, 30-42 (2009).
33. Zhu, Q. et al. Optimal probing of optical contrast of breast lesions of different size located at different depths by US localization. *Technol Cancer Res Treat* **5**, 365-380 (2006).
34. Brooksby, B. et al. Magnetic resonance-guided near-infrared tomography of the breast. *Review of Scientific Instruments* **75**, 5262-5270 (2004).
35. Bevilacqua, F., Berger, A.J., Cerussi, A.E., Jakubowski, D. & Tromberg, B.J. Broadband absorption spectroscopy in turbid media by combined frequency-domain and steady-state methods. *Appl. Optics* **39**, 66498-66507 (2000).
36. Choe, R. et al. Differentiation of benign and malignant breast tumors by in-vivo three-dimensional parallel-plate diffuse optical tomography. *J Biomed Opt* **14**, 024020 (2009).
37. Lee, K. Optical mammography: Diffuse optical imaging of breast cancer. *World J Clin Oncol* **2**, 64-72 (2011).
38. Flexman, M.L. et al. Digital optical tomography system for dynamic breast imaging. *J Biomed Opt* **16**, 076014 (2011).
39. Flexman ML, K.H., Gunther JE, Lim EA, Alvarez MC, Desperito E, Kalinsky K, Hershman DL, Hielscher AH, "Dynamic optical biomarkers for diffuse optical tomography of the breast," submitted to Journal of Biomedical Optics. .
40. Flexman ML, K.H., Gunther JE, Lim EA, Alvarez MC, Desperito E, Kalinsky K, Hershman DL, Hielscher AH, Dynamic optical biomarkers for diffuse optical tomography of the breast," submitted to Journal of Biomedical Optics. *Journal of Biomedical Optics* **18**, 096012-096011-096019 (2013).
41. Thompson, A.M. & Moulder-Thompson, S.L. Neoadjuvant treatment of breast cancer. *Annals of Oncology* **23**, x231-x236 (2012).
42. Rastogi, P. et al. Preoperative Chemotherapy: Updates of National Surgical Adjuvant Breast and Bowel Project Protocols P-18 and B-27. *Journal of Clinical Oncology* **26**, 778-785 (2008).
43. Heil, J. et al. Diagnosis of pathological complete response to neoadjuvant chemotherapy in breast cancer by minimal invasive biopsy techniques. *Br J Cancer* **113**, 1565-1570 (2015).
44. Issa-Nummer, Y. et al. Prospective validation of immunological infiltrate for prediction of response to neoadjuvant chemotherapy in HER2-negative breast cancer -- a substudy of the neoadjuvant GeparQuinto trial. *PLOS ONE* **8**, e79775 (2013).
45. Untch, M. & von Minckwitz, G. Neoadjuvant chemotherapy: early response as a guide for further treatment: clinical, radiological, and biological. *Journal Natl Cancer Inst Monogr* **2011**, 138-141 (2011).
46. Von Minckwitz, G. & Fontanella, C. Selecting the neoadjuvant treatment by molecular subtype: How to maximize the benefit? *The Breast* **22**, S149-S151 (2013).
47. Khokher, S., Mahmood, S., Qureshi, M.U., Khan, S.A. & Chaundhry, N.A. "Initial clinical response" to neoadjuvatn chemotherapy: an in-vivo chemosensitivity test for efficacy in patients with advanced breast cancer. *Asian Pacific Journal of Cancer Prevention* **12**, 939-946 (2011).
48. Enfield, L.C., Gibson, A.P., Hebden, J.C. & Douek, M. Optical tomography of breast cancer – monitoring response to primary medical therapy. *Targeted Oncology* **4**, 219-233 (2009).
49. Chagpar, A.B. et al. Accuracy of Physical Examinaion in Predicting Residual Pathologic Tumor Size in Patients Treated with Neoadjuvant Chemotherapy. *Annals of Surgery* **243**, 257-264 (2006).
50. Yeh, E. et al. Prospective Comparison of Mammography, Sonography, and MRI in Patients Undergoing Neoadjuvant Chemotherapy for Palpable Breast Cancer. *AJR* **184**, 868-877 (2005).

51. Shin, H.J. et al. Comparison of mammography, sonography, MRI, and clinical examination in patients with locally advanced or inflammatory breast cancer who underwent neoadjuvant chemotherapy. *2011* **84**, 612-620 (2011).
52. Keune, J.D. et al. Accuracy of ultrasonography and mammography in predicting pathologic response after neoadjuvant chemotherapy for breast cancer. *The American Journal of Surgery* **199**, 477-484 (2010).
53. Croshaw, R., Shapiro-Wright, H., Svensson, E., Erb, K. & Julian, T. Accuracy of Clinical Examination, Digital Mammogram, Ultrasound, and MRI in Determining Postneoadjuvant Pathologic Tumor Response in Operable Breast Cancer Patients. *Annals of Surgical Oncology* **18**, 3160-3163 (2011).
54. Marinovich, M.L. et al. Meta-Analysis of Magnetic Resonance Imaging in Detecting Residual Breast Cancer After Neoadjuvant Therapy. *J Natl Cancer Inst* **105**, 321-333 (2013).
55. Fang, Q. et al. Combined Optical and X-ray Tomosynthesis Breast Imaging. *Radiology* **258**, 89-97 (2011).
56. Tromberg, B.J. et al. Non-Invasive In Vivo Characterization of Breast Tumors Using Photon Migration Spectroscopy. *Neoplasia* **2**, 26-40 (2000).
57. Pakalniskies, M.G. et al. Tumor Angiogenesis Change Estimated by Using Diffuse Optical Spectroscopic Tomography: Demonstrated Correlation in Women Undergoing Neoadjuvant Chemotherapy for Invasive Breast Cancer? . *Radiology* **259**, 365-374 (2011).
58. Zhu, Q. et al. Breast Cancer: Assessing Response to Neoadjuvant Chemotherapy by Using US-guided Near-Infrared Tomography. *Radiology* **266**, 433-442 (2013).
59. Cerussi, A. et al. Diffuse optical spectroscopic imaging correlates with final pathological response in breast cancer neoadjuvant chemotherapy. *Philosophicals Transactions of the Royal Society* **369**, 4512-4530 (2011).
60. Ueda, S. et al. Baseline Tumor Oxygen Saturation Correlates with a Pathological Complete Response. *Cancer Research* **72**, 4318-4328 (2012).
61. Ross, J.S., Symmans, W.F., Pusztai, L. & Hortobagyi, G.N. in Advances in Clinical Chemistry, Vol. Volume 40 99-125 (Elsevier, 2005).
62. Giancotti, V. Breast cancer markers. *Cancer Letters* **243**, 145-159 (2006).
63. von Minckwitz, G. et al. Definition and Impact of Pathologic Complete Response on Prognosis After Neoadjuvant Chemotherapy in Various Intrinsic Breast Cancer Subtypes. *Journal of Clinical Oncology* (2012).
64. Symmans, W.F. et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J Clin Oncol* **25**, 4414-4422 (2007).
65. Fumagalli, D. et al. A common language in neoadjuvant breast cancer clinical trials: proposals for standard definitions and endpoints. *The Lancet Oncology* **13**, e240-e248 (2012).
66. Masciotti, J.M., Lasker, J.M. & Hielscher, A.H. Digital Lock-In Detection for Discriminating Multiple Modualtion Frequencies with High Accuracy and Computational Efficiency. *IEEE Transactions on Instrumentations and Measurement* **57**, 182-189 (2008).
67. Schneider, P. et al. Fast 3D Near-Infrared Breast Imaging Using Indocyanine Green for Detection and Characterization of Breast Lesions. *Experimentielle Radiologie* **183**, 956-963 (2011).
68. Schreiter, N.F. et al. Optical Imaging of Breast Cancer Using Hemodynamic Changes Induced by Valsalva Maneuver. *Experimentielle Radiologie* **184**, 358-366 (2012).
69. Kim, H.K., Flexman, M., Yamashiro, D.J., Kandel, J.J. & Hielscher, A.H. PDE-constrained multispectral imaging of tissue chromophores with the equation of radiative transfer. *Biomedical Optics Express* **1**, 812-824 (2010).
70. Fawcett, T. An introduction to ROC analysis. *Pattern Recognition Letters* **27**, 861-874 (2006).
71. Zhu, Q. et al. Breast Cancer: Assessing Response to Neoadjuvant Chemotherapy by Using US-guided Near-Infrared Tomography. *Radiology* **266**, 433-442 (2013).
72. Zhu, Q. et al. Utilizing optical tomography with ultrasound localization to image heterogeneous hemoglobin distribution in large breast cancers. *Neoplasia* **7**, 263-270 (2005).
73. Zhu, Q. et al. Noninvasive monitoring of breast cancer during neoadjuvant chemotherapy using optical tomography with ultrasound localization. *Neoplasia* **10**, 1028-1040 (2008).
74. Jiang, S. 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**, 551-560 (2009).
75. Pakalniskis, M.G. et al. Tumor angiogenesis change estimated by using diffuse optical spectroscopic tomography: demonstrated correlation in women undergoing neoadjuvant chemotherapy for invasive breast cancer? *Radiology* **259**, 365-374 (2011).

76. Cerussi, A. et al. Predicting response to breast cancer neoadjuvant chemotherapy using diffuse optical spectroscopy. *Proc Natl Acad Sci U S A* **104**, 4014-4019 (2007).
77. Soliman, H. et al. Functional Imaging Using Diffuse Optical Spectroscopy of Neoadjuvant Chemotherapy Response in Women with Locally Advanced Breast Cancer. *Clinical Cancer Research* **16**, 2605-2614 (2010).
78. Xu, C. et al. Ultrasound-Guided Diffuse Optical Tomography for Predicting and Monitoring Neoadjuvant Chemotherapy of Breast Cancers Recent Progress. *Ultrasonic imaging*, 0161734615580280 (2015).
79. Zhu, Q. et al. Pathologic response prediction to neoadjuvant chemotherapy utilizing pretreatment near-infrared imaging parameters and tumor pathologic criteria. *Breast Cancer Research* **16**, 456 (2014).
80. Jiang, S. et al. Predicting breast tumor response to neoadjuvant chemotherapy with Diffuse Optical Spectroscopic Tomography prior to treatment. *Clinical Cancer Research* **20**, 6006-6015 (2014).
81. Cerussi, A. et al. Predicting response to breast cancer neoadjuvant chemotherapy using diffuse optical spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 4014-4019 (2007).
82. Taroni, P. et al. Breast Tissue Composition and Its Dependence on Demographic Risk Factors for Breast Cancer: Non-Invasive Assessment by Time Domain Diffuse Optical Spectroscopy. *PLoS ONE* **10**, e0128941 (2015).
83. Flexman, M.L. et al. A wireless handheld probe with spectrally constrained evolution strategies for diffuse optical imaging of tissue. *Rev Sci Instrum* **83**, 033108 (2012).
84. Kim, H.K., Gunther, J., Hoi, J. & Hielscher, A.H. A reduced-space basis function neural network method for diffuse optical tomography. *Proc. SPIE* **9319**, 931925 (2015).
85. Ueda, S. et al. Baseline Tumor Oxygen Saturation Correlates with a Pathologic Complete Response in Breast Cancer Patients Undergoing Neoadjuvant Chemotherapy. *Cancer Research* **72**, 4318-4328 (2012).
86. Falou, O. et al. Diffuse optical spectroscopy evaluation of treatment response in women with locally advanced breast cancer receiving neoadjuvant chemotherapy. *Translational oncology* **5**, 238-246 (2012).
87. Schaafsma, B.E. et al. Optical mammography using diffuse optical spectroscopy for monitoring tumor response to neoadjuvant chemotherapy in women with locally advanced breast cancer. *Clinical Cancer Research* **21**, 577-584 (2015).
88. Fauzee, N.J.S., Dong, Z. & Wang, Y.-l. Taxanes: Promising Anti-Cancer Drugs. *Asian Pacific Journal of Cancer Prevention* **12**, 837-851 (2011).
89. Linder, M. & Minckwitz, G.v. in Management of Breast Diseases. (eds. I. Jatoi & M. Kaufmann) 353-372 (Springer Berlin Heidelberg, 2010).
90. Bocci, G., Di Paolo, A. & Danesi, R. The pharmacological bases of the antiangiogenic activity of paclitaxel. *Angiogenesis* **16**, 481-492 (2013).
91. Symmans, W.F. et al. Paclitaxel-induced Apoptosis and Mitotic Arrest Assessed by Serial Fine-Needle Aspiration: Implications for Early Prediction of Breast Cancer Response to Neoadjuvant Treatment. *Clinical Cancer Research* **6**, 4610-4617 (2000).
92. Taghian, A.G. et al. Paclitaxel Decreases the Interstitial Fluid Pressure and Improves Oxygenation in Breast Cancers in Patients Treated with Neoadjuvant Chemotherapy: Clinical Implications. *Journal of Clinical Oncology* **23**, 1951-1961 (2005).
93. Griffon-Etienne, G., Boucher, Y., Brekken, C., Suit, H.D. & Jain, R.K. Taxane-induced Apoptosis Decompresses Blood Vessels and Lowers Interstitial Fluid Pressure in Solid Tumors: clinical Implications. *Cancer Research* **59**, 3776-3782 (1999).
94. Arpino, G. et al. Predictive value of apoptosis, proliferation, HER-2, and topoisomerase IIa for anthracycline chemotherapy in locally advanced breast cancer. *Breast Cancer Res Treat* **92**, 69-75 (2005).
95. Karahaliou, A. et al. Assessing heterogeneity of lesion enhancement kinetics in dynamic contrast-enhanced MRI for breast cancer diagnosis. *The British Journal of Radiology* **83**, 296-306 (2010).
96. Busch, D.R. et al. Optical malignancy parameters for monitoring progression of breast cancer neoadjuvant chemotherapy. *Biomed Opt Express* **4**, 105-121 (2013).
97. Kuo, W.-H. et al. Vascularity Change and Tumor Response to Neoadjuvant Chemotherpy for Advanced Breast Cancer. *Ultrasound in Medicine & Biology* **34**, 857-866 (2008).
98. Mankoff, D.A. et al. Blood Flow and Metabolism in Locally Advanced Breast Cancer: Relationship to Response to Therapy. *The Journal of Nuclear Medicine* **43**, 500-509 (2002).
99. Delille, J.-P., Slanetz, P.J., Yeh, E.D., Kopans, D.B. & Garrido, L. Breast Cancer: Regional Blood Flow and Blood Volume Measured with Magnetic Susceptibility-based MR Imgaing-Initial Results. *Radiology* **223**, 558-565 (2002).

100. Choe, R. et al. Optically Measured Microvascular Blood Flow Contrast of Malignant Breast Tumors. *PLOS ONE* **9**, 1-10 (2014).
101. Madjar, H., Sauerbrei, W., Prömpeler, H.J., Wolfarth, R. & Gulfer, H. Color Doppler and Duplex Flow Analysis for Classification of Breast Lesions. *Gynecologic Oncology* **64**, 392-403 (1997).
102. Corlu, A. et al. Uniqueness and wavelength optimization in continuous-wave multispectral diffuse optical tomography. *Optics Letters* **28**, 2339-2341 (2003).
103. Carpenter, C.M. et al. Monitoring of hemodynamic changes induced in the healthy breast through inspired gas stimuli with MR-guided diffuse optical imaging. *Medical Physics* **37**, 1638-1646 (2010).