Title
ADC as an early indicator of breast cancer response to neoadjuvant treatment

Permalink
https://escholarship.org/uc/item/5vj2714t

Author
Li, Elizabeth

Publication Date
2015

Peer reviewed|Thesis/dissertation

ADC as an early indicator of response to neoadjuvant therapy

by

Elizabeth Li

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Biomedical Imaging

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgement

I would like to thank my advisor Dr. Nola Hylton and Dr. Ella Jones, for exposing me to the field of Breast MR and for their guidance and direction both professionally and personally. I would like to thank Jessica Gibbs, whose attention to detail and constant optimism was indispensable. Thank you to Dr. David Newitt, for being generous with his time and helping me work through IDL issues. I am grateful to Dr. Wen Li and Dr. Lisa Wilmes, for sharing their knowledge and expertise with me, and for many engaging intellectual conversations.

I would like to thank Dr. Roland Krug, Dr. Alastair Martin, and Dr. Viola Rieke for their guidance and time as members of my thesis committee.

Lastly I would like to thank my family and friends for their support as I completed this work.

ADC as an early indicator of breast cancer response to neoadjuvant treatment

Elizabeth Li

Abstract

Quantitative MRI can accelerate drug development by providing non-invasive methods to determine treatment response. The primary aim of this study is to assess the change in normalized apparent diffusion coefficient values (ΔADC_N), derived from diffusion-weighted MRI (DWI), as an alternative method to standard dynamic contrast-enhanced (DCE) MRI for assessing response of primary breast tumors to neoadjuvant chemotherapy. Secondary aims are to: assess the influence of image quality scoring on the predictive performance of ΔADC_N; test correlations between ΔADC_N and change in functional tumor volume (ΔFTV) at early (ΔFTV_2) and late (ΔFTV_4) time points; and assess ΔADC_N of responders versus non-responders.

Methods:

134 patients with primary breast cancers ≥2.5 cm in diameter and high MammaPrint scores were included. 62 and 72 patients received standard and experimental drug regimens respectively. ΔADC_N was determined from DW images acquired at baseline and three weeks into chemotherapy. FTV (70% DCE-MRI enhancement at 2.5 minutes post-contrast) was used as an indication of tumor response throughout treatment. Pathologic complete response (pCR) was determined by histopathology following

surgery. Whole tumor regions of interest (ROIs) and quality scoring was performed on 126 cases, of which 102 had passing quality scores.

Results:

The area under the receiver operating characteristic (ROC) curve (AUC) for ΔADC_N was 0.653 (95% confidence interval (CI) [0.538, 0.768], p=0.00605). The estimated AUC for ΔFTV_2 was not significantly higher than ΔADC_N (mean difference: -0.011±0.086, p=0.896). Using a ΔFTV_4 cutoff of -97.8% as a surrogate endpoint, the AUC estimates were not significantly greater than 0.5.

Image quality did not impact the predictive ability or distribution of ΔADC_N, which increased by 0.836% (95% CI [-0.48, 0.026], p=0.34) with quality scoring. ΔADC_N was not very correlated with ΔFTV_2 or ΔFTV_4. ΔADC_N increased by 9.74% (95% CI [2.24, 17.51], p=0.012) with response in the full cohort.

Summary:

These findings suggest that ΔADC_N may be similar to ΔFTV_2 in predictive performance. While changes in ADC and FTV both reflect changes in tissue properties, they are indicative of independent biological processes. DWI is a promising non-contrast technique that can provide additional information to better predict treatment response.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction:

Quantitative magnetic resonance imaging (MRI) is increasingly used in breast cancer screening, diagnosis and staging. Dynamic contrast-enhanced MRI (DCE-MRI) has unprecedented high sensitivity and resolution to delineate tissue morphology and tumor microvasculature in three dimensions. Comparison of DCE signal enhancement at sequential time points provides information, such as functional tumor volume (FTV), to predict treatment response earlier in treatment than other clinical assessments [1]. A promising technique that is still being explored is diffusion weighted imaging (DWI). DWI signal intensity is derived from thermal or Brownian motion of water and therefore does not require the use of contrast agents. DWI measurements reflect tissue cellularity through apparent diffusion coefficients (ADC). Due to the dense packing of cells in breast tumors, water movement is restricted and ADC values are low relative to normal tissues [2]. In response to chemotherapy, cell density decreases, water diffusivity increases and ADC increases. DCE-MRI and DWI offer complimentary characterizations of breast tumors and are acquired as part of the Investigation of Serial Studies to Predict Your Therapeutic Response With Imaging and Molecular Analysis 2 (I-SPY 2 Trial), a clinical trial investigating the use of MRI and biomarkers to predict response to chemotherapy in breast cancer [3].

I-SPY 2 Trial:

The I-SPY 2 Trial is a multi-center neoadjuvant treatment trial integrating molecular and imaging biomarkers for identifying promising new drugs with a high probability of

success in a subsequent phase III trial (*figure 1*). I-SPY 2 uses an adaptive design that monitors the change in FTV measured by MRI at serial time points during the treatment so that, as the trial proceeds, randomization schema is adjusted accordingly. Drugs are able to "graduate" from I-SPY 2 when the Bayesian predictive probability of achieving 80% success in a subsequent phase III study, but can be dropped for futility if statistical significance is not reached after a predetermined number of patients have been assigned to receive that particular experimental treatment. This minimizes the risk of exposure to therapies deemed unfavorable to patients with certain tumor types, and allows for more targeted phase III trials that include fewer patients [3].

Neoadjuvant chemotherapy not only reduces tumor size to conserve breast tissue, but it also allows for non-invasive monitoring of tumor response for the assessment of new drug efficacy. Not all patients are responsive to systemic chemotherapy, and they may be unnecessarily exposed to toxic treatment without clear clinical benefits. Targeted therapies based on gene expression and receptor status may be more effective and are increasingly being used in the customization of patient care. For example, luminal tumors are responsive to hormone therapy, whereas human epidermal growth factor receptor 2 positive (HER2+) tumors can be treated with a myriad of anti-HER2 therapies such as Herceptin (Trastuzumab). Non-invasive imaging with MRI provides information to assess the efficacy of new targeted therapies and may quickly identify non-responders for alternative treatment. In I-SPY 2, taxane-based treatment (Paclitaxel) is used in combination with an experimental targeted drug, followed by anthracycline-based treatment (doxorubicin and cyclophosphamide, AC). Patients who

are HER2+ will also receive Herceptin in addition to Paclitaxel in the first set of the treatment regimen (*figure 1*).
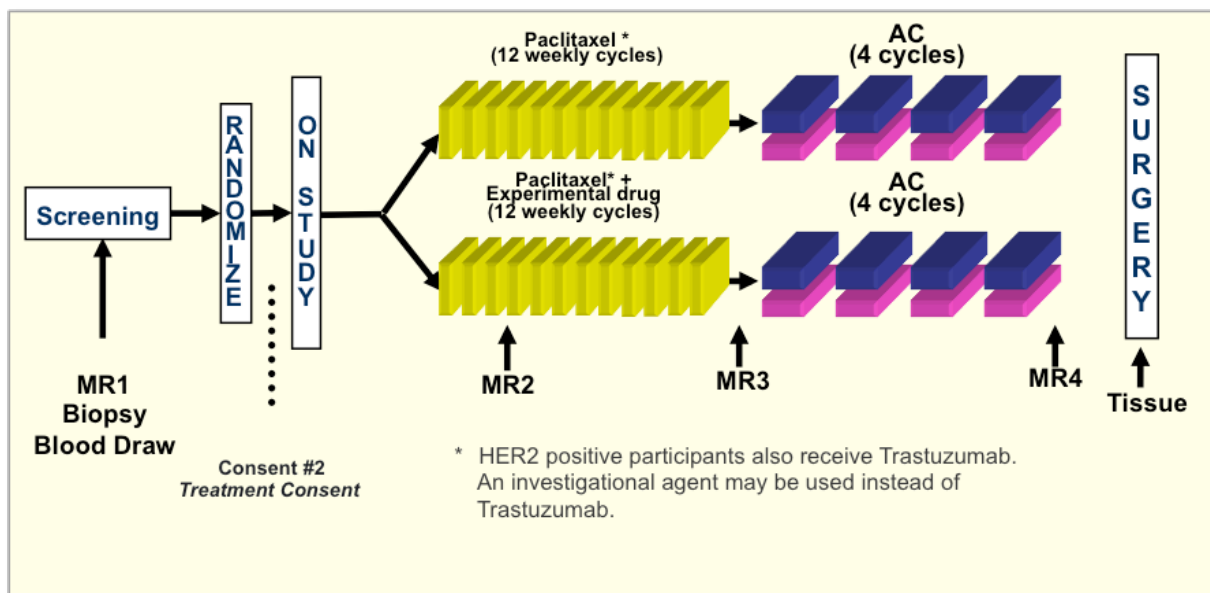

Figure 1: I-SPY 2 simplified study schema. A taxane-based compound (Paclitaxel) was administered sequentially with a combination of doxorubicin and cyclophosphamide (AC), anthracycline-based chemotherapy. Randomization included factors such as MRI tumor volume and molecular biomarkers. Note: each AC cycle spanned 21 days.

Participants in the I-SPY 2 Trial undergo sequential DCE- and DW-MRI exams during the course of treatment (*figure 1*). MR data is acquired at baseline (MR1), following 3 weeks of chemotherapy (MR2), after the completion of the first regimen (MR3), and before surgery (MR4). Functional tumor volume (FTV), calculated based on DCE-MRI, is one of the biomarkers being evaluated in I-SPY 2 for prediction of treatment response. FTV is computed as the sum of all voxels meeting a minimum threshold for initial percent enhancement (PE) defined as PE (PE = 100 x $(S_1 - S_0)/S_0$), where S_0 and S_1 represent the signal intensities of each voxel in the pre-contrast and early post-contrast images respectively. Early change in FTV is measured from baseline

at MR1 to MR2 (ΔFTV_2), while late change in FTV is measured from MR1 to MR4 (ΔFTV_4).

Pathologic complete response (pCR), determined histologically by the absence of residual tumor at the time of surgery, is used as a surrogate endpoint for response to therapy. As the collection of long-term survival data for I-SPY 2 participants is still underway, the model for pCR outcome with FTV as a predictor is used. Although the primary goal of each I-SPY 2 MRI exam is to obtain DCE-MRI images for FTV determination, the complementary DWI measurement can be added to each scan at no additional cost and with no added inconvenience or risk to the patient. In fact, a sub-study funded by the American College of Radiology Imaging Network (ACRIN) 6698 is being conducted under I-SPY 2 to determine whether the early change in ADC values is predictive of response.

In 2013, a poly ADP ribose polymerase (PARP) inhibitor graduated from the trial. While the mechanism of PARP inhibition is not fully understood, PARP inhibitors have been shown to potentiate the cytotoxic activity of DNA alkylation by trapping PARP at sites of DNA damage [4]. Inhibitors of DNA repair enzymes such as PARP-1, PARP-2, BRCA-1 and BRCA-2 decrease the rate at which repair processes such as base excision repair, homologous recombination, and non-homologous end joining can occur. Another small-molecule dual-inhibitor of HER2 and epidermal growth factor receptor (EGFR) kinases also graduated from the trial in 2014. HER2-positive patients have a 2- to 20- fold amplification of the HER2/neu oncogene that is responsible for HER2 production, and can be found in 30% of breast tumors [5]. Tumors with an increased

rate of DNA repair, or an overabundance of receptors that are associated with increased cell growth are susceptible to chemotherapy. Suppression of repair enzymes activities and inhibition of receptors such as HER2 are important strategies in the development of anti-cancer therapies.

Diffusion Weighted Imaging:

Various breast tumor types, both mass-like and diffuse, have been investigated and characterized using DWI [2,6-7]. The benefits of DWI in the breast have been observed clinically. In a study by Partridge *et al.*, the complementary information provided by DWI in addition to DCE-MRI resulted in a 10% overall increase in the positive predictive value (PPV) with an ADC threshold set for maximum specificity, and a 17% increase in PPV for lesions 1 cm or smaller [8]. This increased diagnostic power can be especially useful in more complicated cases and earlier during the course of disease, where pCR after neoadjuvant chemotherapy can be most predictive of long-term survival [9].

The diffusion-weighted signal is acquired through T2-weighted single-shot spin-echo echo-planar sequences (SS-EPI) with diffusion-sensitizing gradients. The diffusion-sensitizing gradients act as exponential weighting factors applied to a T2-weighted image, and are represented mathematically by b-values. Diffusion weighted images are derived using a minimum of four separate image acquisitions: three high b-value images and one low b-value image S_0. The diffusion-weighted image S_{ij} is determined using equation *(1)* for the *i*th b-value and each orthogonal gradient direction *j*:

$$S_{ij} = S_0 e^{-b_i ADC_{ij}} \qquad (1)$$

Clinically, high-end b-values range from 600 to 1500 s/mm^2 [10]. The b-value is given by equation *(2)*,

$$b_i = \gamma^2 G_i^2 \delta_i^2 \left(TR_i - \frac{\delta_i}{3}\right) \qquad (2)$$

where γ is the gyromagnetic ratio, G is the gradient strength, δ is the gradient duration and TR is the repetition time. It can be seen from equation *(2)* that increasing δ or G will most efficiently increase the resulting b-value, ultimately maximizing the sensitivity to smaller changes in diffusion. However, higher b-values result in lower signal intensity due to increased dephasing with diffusion. The signal intensity from a T2-weighted image lacking any weighting by motion-sensitizing gradients has a b-value of zero s/mm^2, and is typically the low b-value, or b_0 image S_0. $\bar{S}$ is the rotationally invariant, geometric mean signal intensity of the three gradient directions acquired at the same b-value. $\bar{S}$ can be used to determine *in vivo* diffusion constants termed apparent diffusion coefficient (ADC) values on a pixel-by-pixel basis. Collectively these values form an ADC map and can be determined using one high b-value b_i and one low b-value b_0 as shown in equation *(3):*

$$ADC = -\ln\left(\frac{\bar{S}_i}{S_0}\right)/(b_i - b_0) \qquad (3)$$

When ADC values are derived from more than two b-values such as 0, 600, 800, and 1000 s/mm^2, a log-linear least-squares regression *(4):*

$$\ln(S_i) = \ln(S_0) - b_i \cdot ADC \qquad (4)$$

is performed in order to utilize all available information. Though Mukherjee *et al.* argued that the use of more than two b-values might be redundant in clinical application [10],

this refers to MR imaging of the human brain and has not been the case for breast tissue. However, the benefits of multiple b-value acquisitions have not been observed over traditional, two b-value DWI to date in breast imaging. The incorporation of more b-values would increase the samples obtained, theoretically increasing the accuracy of the ADC values obtained [11]. Because *in vivo* diffusion in breast tissue is more complicated than that in the brain, standardization of DWI in breast tissue is needed for future clinical use [7,11].

Previous work on the changes in ADC values (ΔADC) throughout the course of neoadjuvant chemotherapy have shown that positive ΔADC is associated with treatment response [12-14]. Additionally, the benefits of using normalized ΔADC (ΔADC_N, where ADC_N is the ratio of tumor ADC to ADC of normal fibroglandular tissue) rather than absolute ΔADC values (ΔADC_A) have been demonstrated [15]. Early ΔADC_N may be predictive of treatment response, and existing I-SPY 2 data will be used to further investigate these findings.

For this study, MR images from patients in the experimental and standard of care cohorts were analyzed. Whole tumor ADC values were determined at baseline (MR1) and early in treatment (MR2). Patients' pCR status was the primary outcome.

Specific Aims:

Primary Aim:

- To determine whether ΔADC_N is predictive of pCR or ΔFTV_4 (an alternative endpoint defined as $Response_{FTV}$).

Secondary Aims:

- To assess the influence of image quality scoring on the predictive performance of ΔADC_N.

- To determine if there is a correlation between ΔADC_N and ΔFTV at both early (ΔFTV_2) and late (ΔFTV_4) time points.

- To assess the difference in mean ΔADC_N values between responders and non-responders with pCR and $Response_{FTV}$ as outcomes.

Materials and Methods:

Patients with biopsy-confirmed invasive breast cancer of ≥2.5 cm in diameter, with no prior chemotherapy or radiation therapy and high MammaPrint scores were enrolled. Their biomarker signatures were hormone receptor positive (estrogen or progesterone ER+/PR+) and/or HER2+ or triple negative (ER-, PR-, and HER2-). All patients had given their written informed consent to participate in I-SPY 2. This trial has been open to enrollment since March 2010 and has a projected completion date of September 2017. Patient data used for the present work were acquired between April 2010 and January 2013. FTV measurements were performed at participating sites or at the UCSF Imaging Core Lab using a semi-automated software system (Aegis 4D Visualization Software, Hologic Incorporated).

Of 134 patients from nineteen study locations, 62 received standard neoadjuvant chemotherapy and 72 received the experimental drug in addition to the taxane-based treatment during the first set of neoadjuvant chemotherapy.

In ACRIN 6657, the imaging component of I-SPY 1 Trial, Hylton *et al.* found that ΔFTV is a stronger predictor of pCR than clinical assessment [1]. A 50% decrease in tumor volume has been associated with significant differences in recurrence free survival [16]. Because long-term survival information is not yet available for I-SPY 2 patients, ΔFTV$_4$ was used as an alternative surrogate endpoint in this study.

Imaging protocol:

Imaging was performed with patients in the prone position on either 1.5T or 3.0T whole body MRI scanners using dedicated breast coils. As part of the image quality assurance and control (QA/QC) requirements in I-SPY 2, sequential imaging exams are performed using the same field strength and same scanner model used in the baseline visit. This minimizes image quality variations between and within patients.

DWI was acquired using a 2D diffusion-weighted SS-EPI technique. Bilateral axial images were acquired with minimum echo time (TE), repetition time (TR) of at least 4,000 ms, a field of view (FOV) in the range of 26-36 cm and a matrix size of 128-192 in both frequency and phase directions. A minimum of 2 averages were used and a parallel imaging factor of at least 2 was used to achieve scan durations of approximately four minutes. A minimum of two b-values were used: 0 s/mm^2 and either 600 or 800 s/mm^2. The slice thickness was 3-5 mm and the desired in-plane resolution was less than 1.9 mm. Active fat saturation techniques were implemented. DWI acquisitions were completed prior to any contrast-enhanced sequences.

DCE T1-weighted images were obtained using an axial 3D gradient echo (GE)

sequence with minimum TE, TR of 4-10 ms, flip angle of 10-20 degrees, and a maximum of two repetitions. FOV was between 26-36 cm in both frequency and phase direction, and slice thickness was ≤2.5 mm. FOV and slice thickness were set to achieve complete bilateral coverage for each exam. The acquisition matrix was 384-512 by at least 256 for a maximum in-plane resolution of 1.4 mm. At least 60 slices were acquired. Scan time was between 80 and 100 seconds per time point. One pre-contrast T1 image was acquired and checked for proper fat saturation, and post-contrast T1 images were obtained for at least 8 minutes after the injection of an FDA-approved gadolinium-based contrast agent. All patients were given the same contrast agent brand at sequential visits. Injection was administered at a rate of 2 ml/s at a concentration of 0.1 mmol/kg body weight and was followed by a 20 ml saline flush. Gain settings were held constant for pre- and post-contrast T1 images. Active fat saturation methods were recommended but were left at the discretion of each site.

Processing of DWI data:

An in-house software tool using the environment of Interactive Data Language (IDL, Research Systems) was used for image analysis. ADC maps (*figure 2B*) were generated using DWI data for MR1 and MR2. When two b-values were acquired, equation *(3)* was used to produce the ADC map. When greater than two b-values were available, least-squares regression for the natural log of the signal versus b-value was determined, which resulted in a slope of negative ADC *(4)*. The post-contrast DCE subtraction images (*figure 2C*) were used to determine the location of the tumor. Whole-

tumor regions of interest (ROIs) were manually drawn on the ADC maps to include enhancing regions on the subtraction image (*figure 2C*) and DWI images (*figure 2A*) while isolating the same darker regions on the ADC maps. In general, higher b-value DWI images were used for determining areas of lowest diffusivity for ROI inclusion. ROIs were delineated to include as much solid tumor as possible on contiguous slices while avoiding fat and non-malignant fibroglandular tissue to eliminate partial volume averaging effects. Regions of necrosis and susceptibility artifacts from biopsy clips were excluded.
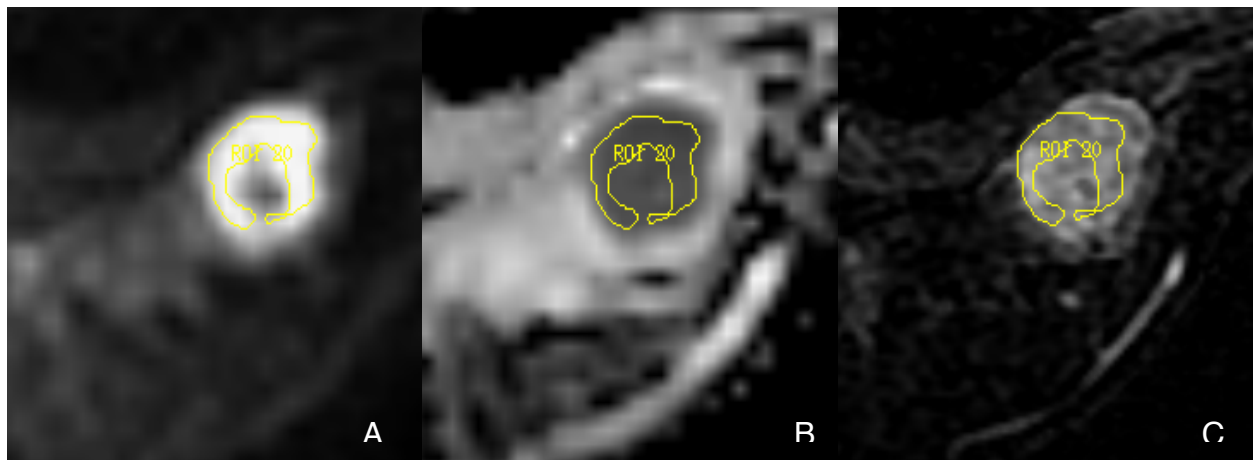

Figure 2: Diffusion processing example. DWI b=800 s/mm^2 (A), ADC map (B), DCE subtraction image (C). Whole-tumor ROIs were drawn manually on (B), while taking into account (A) and (C) to avoid areas of low diffusivity, susceptibility artifacts, and non-enhancing areas. Edges of tumors were also disregarded as to avoid partial-volume averaging.

A baseline level of enhancement from normal fibroglandular tissue was determined to account for normal variability between patients [11]. Normal tissue ROIs were delineated in an area of normal-appearing tissue in the contralateral breast on the ADC map (*figure 3B*). DWI images and pre-contrast images were taken into account to exclude areas of fat, since DWI slices were roughly twice as thick as DCE images.

Mean ADC values were calculated for all normal tissues.   In order to obtain a representative sample of normal fibroglandular tissue, circular ROIs with a diameter of at least 6.0 mm were drawn on a minimum of four contiguous ADC map slices. ROIs were drawn in similar anatomical regions on the contralateral breast whenever image quality was sufficient.
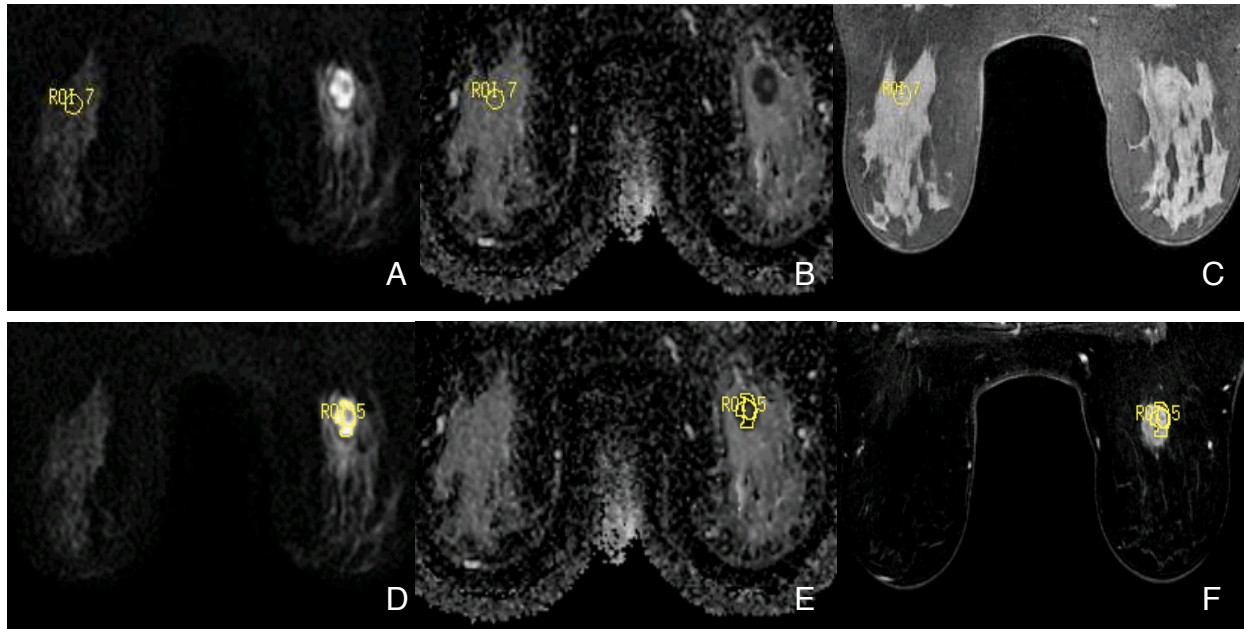

Figure 3: Ideal case for ROI delineation. The high b-value diffusion image series, where b=800 s/mm^2 (A, D), was taken into account as ROI delineation occurred on the ADC map (B, E). The pre-contrast DCE image (C) was referenced to determine optimal normal tissue ROI placement, which was overlaid in a similar anatomical location in the contralateral breast. Whole tumor ROIs were drawn with regard to a subtraction image (F) created using two early time points of the DCE series to better visualize enhancing tissue.

Mean ADC of each malignant lesion was divided by the corresponding mean ADC of fibroglandular tissue to obtain normalized ADC values ($ADC_N = ADC_{Tumor} / ADC_{Fibroglandular}$). ADC_N for each visit was then used to determine the percent change in normalized ADC between visits ($\Delta ADC_N = 100 \times [(ADC_{N\_MR2} / ADC_{N\_MR1}) - 1]$). ΔADC_N was used for analyses unless stated otherwise.

An image quality scoring system was implemented to differentiate high quality images from lower quality images. The scoring system evaluated the quality of fat suppression (*figure 4A*), the presence of artifacts (displacement (*figure 4B,C*), ghosts, or warping), and signal-to-noise ratio (SNR) as these impact lesion conspicuity and ROI set confidence, which was also scored. This scoring system was used to provide a
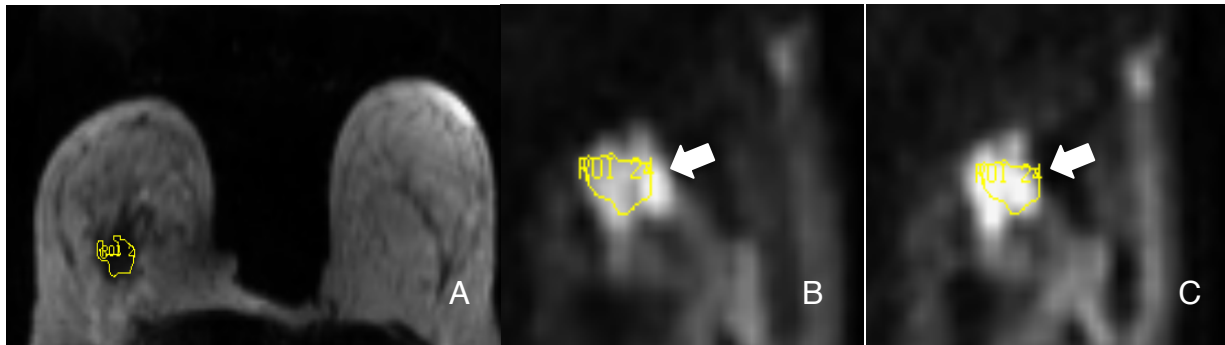

Figure 4: Common DWI artifacts. Failure of fat suppression techniques can be seen in the diffusion image (A). Here, breast density is also low and very little fibroglandular tissue is observed. Patient motion between acquisitions (B, C) was common. (B) and (C) depict the same ROI overlaid on b=800 s/mm^2 diffusion images from two orthogonal gradient directions. The final diffusion image is the geometric mean of the three gradient directions as described by equation *(1)*.

quality threshold for inclusion in analyses (*figure 5*). Quality scores for fat saturation, appearance of artifacts, and SNR were assigned separately, though double penalization was avoided. ROI confidence was scored per exam. Individual visits and ROIs were scored separately and combined to determine whether a patient was included in the final analyses.
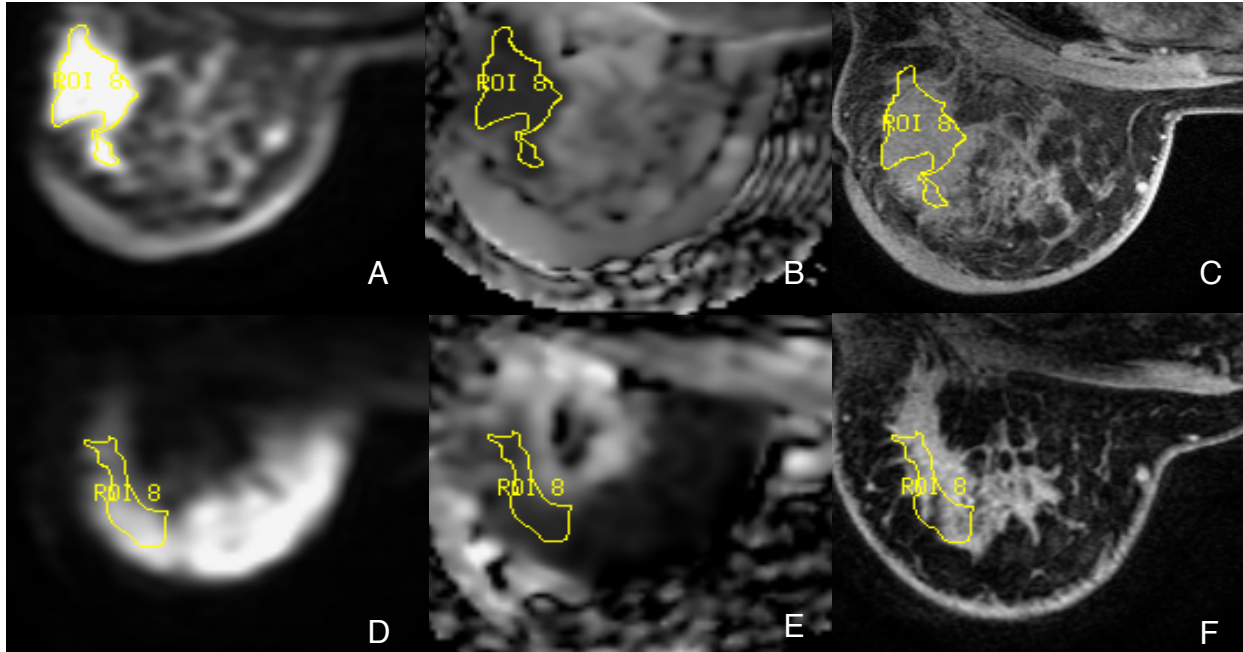
Figure 5: Quality scoring example. Bilateral view b= 800 s/mm^2 DW image (A), ADC map (B), and subtraction image(C) from MR1 and ipsilateral view of b= 800 s/mm^2 DW image (D), ADC map (E), and subtraction image (F) from MR2 for a single patient. The quality scoring results of (A) and (D) were used to assess inclusion of ΔADC_N obtained from the corresponding ADC maps (B) and (E). To assess the appearance of ghosting, phase artifacts, motion between b-value acquisitions (*figure 4B, C)* or warping, all b-value images were considered. To assess the efficacy of fat saturation (*figure 4A*) and signal to noise ratio (SNR), the b_0 image was compared to the pre-contrast DCE-MRI (*figure 3C*). As a quality-scoring example, (A) was of reasonable quality in terms of fat suppression, artifacts, and SNR. (D) had poor fat saturation and extensive artifacts but had reasonable SNR. Confidence in the ROI sets for MR1 and MR2 were medium and low respectively. Though the quality of MR1 was reasonable, the quality of MR2 and low confidence in the ROI set resulted in the exclusion of this case.

Processing of DCE data:

Contrast enhancement was calculated on a voxel-by-voxel basis (*figure 6*) by using the signal enhancement ratio (SER) method, which compares early and late percent enhancement (PE; SER = PE_{early}/PE_{late}) as a measure of tumor vascularity and therefore function [17]. FTV was determined by summing voxel volumes meeting a threshold of

70% PE at 2.5 minutes post-contrast injection over baseline [16,18]. Change in FTV (ΔFTV) was determined in a similar manner to ΔADC_N ($\Delta FTV = 100 \times [(FTV_{MRn} / FTV_{MR1}) - 1]$), where n is equal to 2 for early ΔFTV (ΔFTV_2) or is equal to 4 for late ΔFTV (ΔFTV_4).

Statistical Methods:

The full cohort was analyzed in subsets of responders and non-responders. Subsets of patients with and without pCR receiving standard of care or experimental drug treatments were also analyzed.

A test of normality was performed using the Shapiro-Wilk method. Non-parameterized ROC

Figure 6: Volumetric analysis. DCE post-contrast 2.5 minutes (A), tumor volume calculations that meet at least 70% enhancement (B). Percent enhancement was defined as PE = $100 \times (S_1 - S_0)/S_0$, where S_0 and S_1 were the signal intensities of each voxel in the pre-contrast and 2.5 minutes post-contrast images respectively.

analysis was used to explore ΔADC_N thresholds that maximized both sensitivity and specificity using 2000 bootstrap replicates with pCR as the outcome [19]. ROC analysis based on DeLong's method [20] was used to compare AUC curves of ΔFTV_2 and ΔADC_N as predictors of pCR. A threshold ΔADC_N of 18.18% and a ΔFTV_2 cutoff of -68.77% were established through ROC analysis for maximum sensitivity and specificity. The optimal threshold for ΔFTV_4 as an alternative outcome within these subsets was also analyzed and was defined as Response$_{FTV}$. A ΔFTV_4 cutoff of -97.80% was
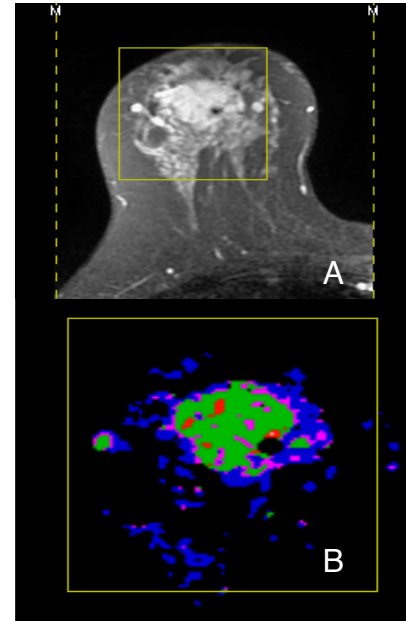
determined through ROC analysis for maximum specificity and sensitivity. Since FTV is expected to decrease over the course of treatment, ΔFTV_4 of less than this value was defined as a responder. ROC analysis was repeated using Response$_{FTV}$ as the outcome. A threshold ΔADC_N of 9.652% was established based on maximizing sensitivity and specificity of ROC analysis with Response$_{FTV}$ as a surrogate endpoint.

In order to ensure image quality scoring standards were appropriate, density plots were determined for the subset used for analyses as well as all fully processed scans regardless of quality.

Correlation between ΔADC_N and ΔFTV_2 as well as ΔFTV_4 was calculated using Spearman's rank correlation (ρ), or Kendall's rank correlation (τ_b) when ties occurred in the dataset. Statistical analyses were performed using ΔADC_N as a predictor. Absolute ΔADC values (ΔADC_A) were also assessed in correlations with ΔFTV.

To test the significance of the difference in ΔADC_N between responders and non-responders in regards to both pCR and Response$_{FTV}$, a two-tailed Mann-Whitney *U* test was performed. The sensitivity and specificity of significant shifts in ΔADC_N were extracted from the corresponding ROC curve.

All statistical analysis was performed using R (R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/). 95% confidence intervals were used. A p value of less than or equal to 0.05 was considered statistically significant.

Results and discussion:

Study population:

Of the 134 patients enrolled, 2 patients did not complete one or more of the relevant exams, 6 patients were excluded due to corrupted data or poor image quality. Once DWI quality scoring was employed, 24 cases failed to meet the quality standards, including 1 case with no contralateral breast image (*table 1*). The final cohort of 102 patients had reasonable image quality at MR1 and MR2.

Enrollment (n = 134)
• 2 Patients withdrew from study

DWI Processing (n = 132)
• 6 Unable to analyze further

Quality Assessment (n = 126)
• 24 Failed quality scoring
 • 1 No contralateral

Statistical Analysis (n = 102)
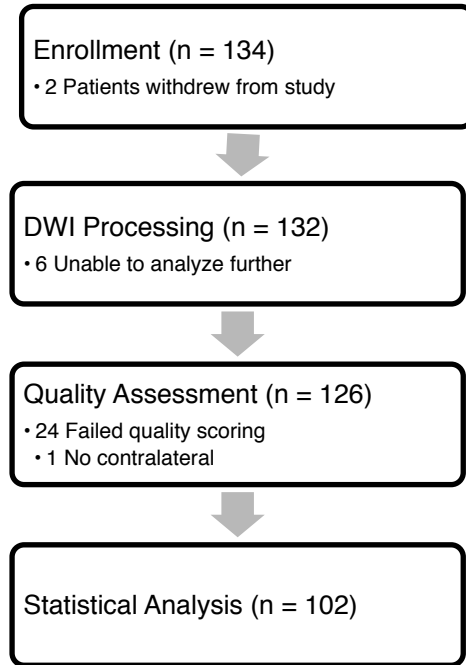
Table 1: Study workflow. Reasons for exclusions were based on availability of exam and image quality.

Patient characteristics:

Patients' tumor characteristics are listed in table 2. Their hormone receptor (HR) and HER2 statuses as well as risk of recurrence [21-22] as reflected by ultra high MammaPrint scores [23], were recorded (*table 2*). In the full cohort, 49 (48.04%) were HR+, 14 (13.73%) were HER2+, 16 (15.69%) were triple negative, and 43 (42.16%) had ultra high MammaPrint scores.

	Full cohort (n=102) (%)	Standard (n=41) (%)	Experimental (n=61) (%)
HR+	49 (48.04)	23 (56.10)	26 (42.62)
HER2+	14 (13.73)	14 (34.15)	-
Triple Negative	16 (15.69)	9 (21.95)	7 (11.48)
Ultra High MP	43 (42.16)	6 (14.63)	37 (60.66)
pCR	34 (33.33)	8 (19.51)	26 (42.62)
non-pCR	68 (66.67)	33 (80.49)	35 (57.38)

Table 2: Patient characteristics. Enrolled individuals and a variety of subsets included in analysis are displayed. * Missing data includes patients with exams not included in analysis.

Summary statistics:

Four combinations of b-values were used for DWI acquisition. 95 (93%) of the patients included in the analysis were the result of using b-values of 0 and 800 s/mm^2 for both MR1 and MR2 exams. Other combinations included: 0 and 600 s/mm^2 (2 cases); 0, 100, 600, and 800 s/mm^2 (3 cases); 0, 100, 600, 800, and 1000 s/mm^2 (9 cases).

Summary statistics of ΔADC_N using pCR as an outcome:

Across the full cohort the mean ΔADC_N was 18.03 ± 19.81%. Mean ΔADC_N was 25.71 ± 22.81% for patients that exhibited pCR and was 14.19 ± 17.04% for non-pCR patients from the full cohort (*figure 7*).

Patients that received the standard treatment exhibited a mean ΔADC_N of 12.66 ± 14.86%. The mean ΔADC_N was 20.51 ± 16.01% for pCR and was 10.75 ± 14.30% non-pCR standard treatment groups respectively. Patients that received the experimental drug exhibited a mean ΔADC_N of 21.65 ± 21.87%. Mean ΔADC_N was 27.31 ± 24.57%

with pCR and was 17.88 ± 21.45% for non-pCR for patients of the experimental cohort.

Boxplot summary statistics for groups involved in analyses are in tables 3A and 3B.
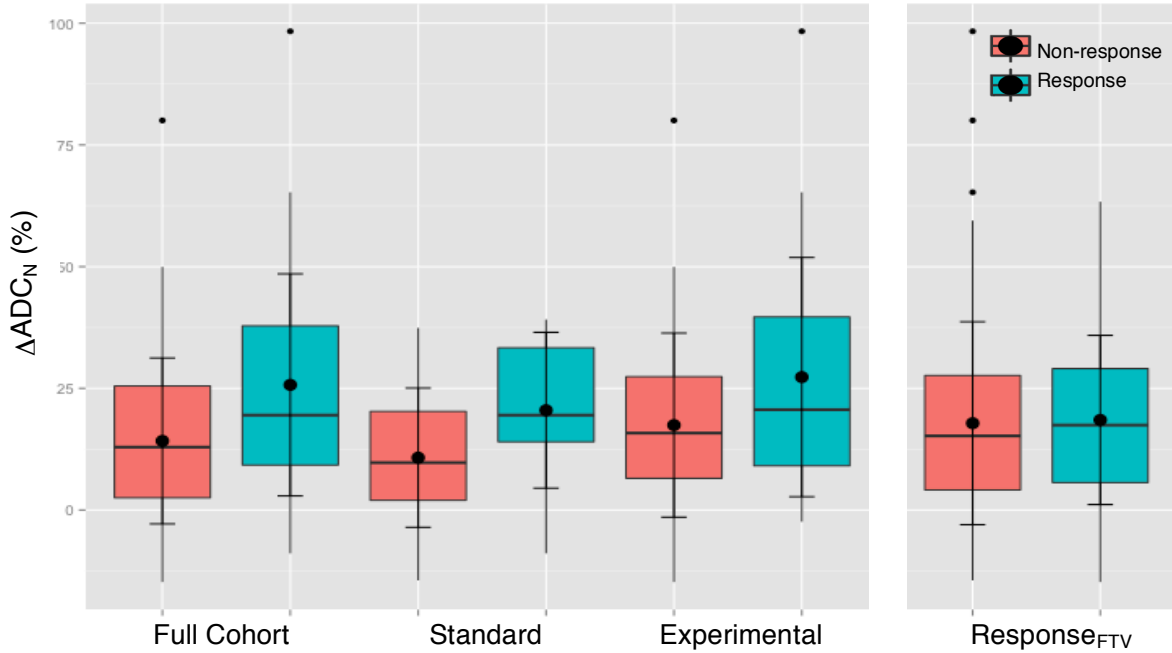

Figure 7: Boxplots of ΔADC$_N$ by response and treatment type. Shown in blue are responders and in red, non-responders. Corresponding mean and median for each cohort are listed in table 3.

Table 3A	Full Cohort (n=102)	Standard (n=41)	Experimental (n=61)	Response$_{FTV}$ (n=30)	Non-Response$_{FTV}$ (n=72)
Mean	18.03	12.66	21.66	18.50	17.84
Median	15.55	11.19	17.39	17.44	15.23

Table 3B	Full Cohort (n=102)		Standard (n=41)		Experimental (n=61)	
	pCR	Non-pCR	pCR	Non-pCR	pCR	Non-pCR
Mean	25.71	14.19	20.51	10.75	27.31	17.44
Median	19.49	12.91	19.49	9.73	20.63	15.81

Table 3A/B: Summary statistics. Boxplot summary statistics of ΔADC$_N$ for groups included in analyses (A) and by presence of pCR (B). Response$_{FTV}$ and non-Response$_{FTV}$ groups statistics were included in table 3A.

Summary statistics of ΔFTV and Response$_{FTV}$ as an outcome:

For the full cohort, mean ΔFTV_2 was -33.82 ± 60.11% and -87.44 ± 17.91% respectively. For all patients with pCR, mean ΔFTV_2 was -43.87 ± 70.81% and for non-pCR, it was -28.78 ± 51.66%. For patients who received the standard of care, mean ΔFTV_2 was -25.65 ± 60.71%. Mean ΔFTV_2 of the standard treatment group was -45.61 ± 41.99% when pCR was exhibited and was -20.81 ± 64.01% without pCR. For patients of the experimental drug group, mean ΔFTV_2 was -39.32 ± 59.57%. Mean ΔFTV_2 for patients who received the experimental treatment was -43.41 ± 82.16% with pCR and was -36.29 ± 35.82% without pCR.

Mean ΔFTV_4 was -87.44 ± 17.91%, -91.08 ± 18.47, and -85.53 ± 17.46% for the full cohort, for all pCR patients and for all non-pCR patients respectively. Mean ΔFTV_4 of patients who received the standard treatment was -85.36 ± 18.17%. For patients of the standard of care group, mean ΔFTV_4 was -90.13 ± 17.93% with pCR and was -84.08 ±18.33% without pCR. For all patients that received the experimental treatment, mean ΔFTV_4 was -88.80 ± 17.76%. Mean ΔFTV_4 was -91.38 ± 19.00% and was -86.85 ± 16.80% for patients of the experimental arm that exhibited and did not exhibit pCR respectively.

ROC analysis using ΔFTV_2 as a predictor of pCR resulted in an AUC estimate of 0.664 (95% CI [0.541, 0.787], p=0.00353) at a ΔFTV_2 threshold of -68.77% based on the maximization of specificity and sensitivity. From the ROC curve using ΔFTV_4 for the prediction of pCR (AUC 0.696, 95% CI [0.578, 0.814], p=0.000846), a threshold ΔFTV_4 of -97.80% was used as the cutoff for the alternative endpoint, Response$_{FTV}$. Using this

premise for the full cohort, mean ΔADC_N was 18.50 ± 17.38% for responders and 17.84 ± 20.85% for non-responders (*table 3A*). A boxplot with Response$_{FTV}$ as an outcome was also included in figure 7.

Primary Aim: ΔADC_N and prediction of pCR and Response$_{FTV}$

The Shapiro-Wilk test was used to test the sample distribution for non-normality prior to further analysis (p < 0.00001). Non-parameterized ROC analysis of ΔADC_N for the full cohort resulted in an AUC estimate of 0.653 (95% CI, [0.538, 0.768], p=0.00605) (*table 4*) with a specificity of 0.662 and a sensitivity of 0.618 (*figure 8A*). Using the AUC estimates for ΔADC_N and ΔFTV_2, DeLong's test to compare non-parameterized ROC curves revealed that the difference between the ROC curves for ΔADC_N and for ΔFTV_2 was not significant, with a mean AUC difference of -0.011 ± 0.086 (p=0.896).

When interrogating the standard (*figure 8B*) and experimental (*figure 8C*) therapy groups, AUC estimates were not significantly greater than 0.5 since the 95% confidence intervals included 0.5 (*table 4*).

Using a ΔFTV_4 cutoff of -97.80% as an alternative outcome, (Response$_{FTV}$, *figure 8D*), the AUC was estimated at 0.533 (95% CI [0.411, 0.656], p=0.30). This resulted in a maximum specificity of 0.698 and maximum sensitivity of 0.424, at a threshold ΔADC_N of 9.652%. AUC estimates were also determined for experimental and standard treatment groups using Response$_{FTV}$, but were less than 0.5 (*table 4*).

AUC for ΔADC_N by response metric		
AUC ([95% CI], p)	pCR	Response$_{FTV}$
Full Cohort (n=102)	0.653 ([0.538, 0.768], 0.00605)	0.533 ([0.411, 0.656], 0.300)
Experimental (n=61)	0.614 ([0.468, 0.760], 0.0659)	0.497 ([0.329, 0.666], 0.574)
Standard (n=41)	0.689 ([0.464, 0.915], 0.0521)	0.441 ([0.202, 0.681], 0.188)

Table 4: AUC estimates for ΔADC_N by response metric: pathologic complete response (pCR) or response based on threshold ΔFTV_4 (Response$_{FTV}$). 95% confidence interval was used, and groups were separated by treatment type.
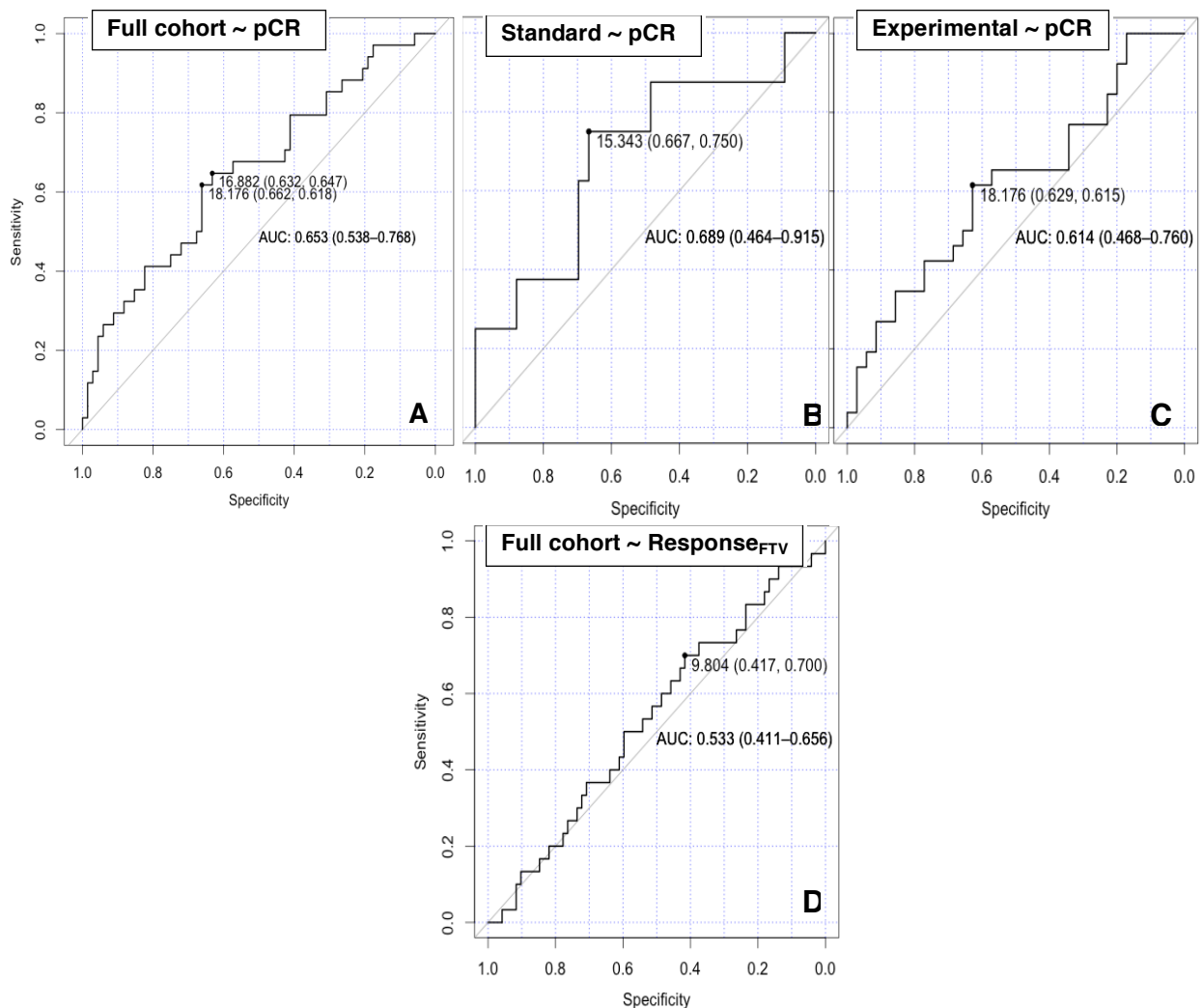

Figure 8: ROC curves of ΔADC_N and prediction of response. AUC estimates were plotted for the full cohort (A), patients who received the standard (B) and experimental (C) therapies using pCR as an outcome and with Response$_{FTV}$ was used as an outcome (D).

Secondary Aim 1: Influence of quality scoring on ΔADC$_N$ for prediction of pCR

The distributions of ΔADC$_N$ of processed cases with passing quality scores (n=102), failing quality scores (n=23) and of all processed cases regardless of quality scores (n=126) were compared (*figure 9A*). The distribution of ΔADC$_N$ with and without quality control remained unchanged and resulted in an estimated mean difference of 0.836% (95% CI [-0.48, 0.026], p=0.34). However, there was greater variation in ΔADC$_N$ in the cases excluded due to failing quality (n=23), with a standard deviation of ± 30.62%
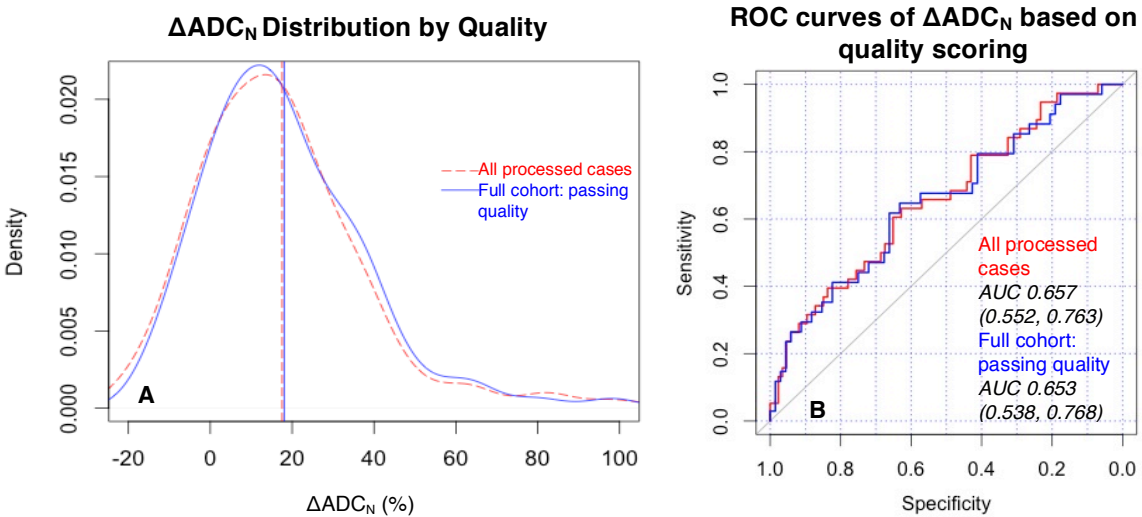

Figure 9: Density curve of ΔADC$_N$. (A) and ROC curves based on quality scoring (B). Quality scoring (A, blue) resulted in a similar distribution of ΔADC$_N$ to the distribution when all processed cases are included (B, red). Quality scoring had little impact on the AUC estimate (B, blue) as compared to the full cohort prior to quality scoring implementation (B, red).

compared to a standard deviation of ± 19.81% from cases with passing image quality (n=102). ROC curves with and without quality scoring were compared by DeLong's test which resulted in a mean difference and standard deviation of -0.0042 ± 0.079 (p=0.958) (*figure 9B*) [20].

Secondary Aim 2: Correlation of ΔADC_N and ΔFTV

For the full cohort, there were no significant correlations between ΔADC_N and ΔFTV_2 for all cases ($\rho=-0.12$, $p=0.23$), for those with pCR ($\rho=0.03$, $p=0.87$), or for cases without pCR ($\rho=-0.11$, $p=0.37$). In patients that received the standard treatment, ΔADC_N was not correlated with ΔFTV_2 ($\rho=0.12$, $p=0.43$). There was a statistically significant positive correlation between ΔFTV_2 and ΔADC_N in the standard treatment group that exhibited pCR ($\rho=0.81$, 95% CI [0.22, 1.00], $p=0.022$, n=8). For patients that received standard treatment who did not exhibit pCR, ΔADC_N was not correlated with ΔFTV_2 ($\rho=0.03$, $p=0.89$). In patients who received the experimental drug, a negative correlation of ΔADC_N and ΔFTV_2 was trending toward significance ($\rho=-0.25$, 95% CI [-0.47, -0.01], $p=0.054$). ΔADC_N for patients of the experimental drug group did not correlate with ΔFTV_2 with ($\rho=-0.14$, $p=0.49$) or without ($\rho=-0.25$, $p=0.14$) pCR.

The correlation between ΔADC_N and ΔFTV_4 was not significant for the full cohort ($\rho=-0.02$, $p=0.80$), the full cohort with pCR ($\rho=0.11$, $p=0.39$), or the full cohort without pCR ($\rho=0.01$, $p=0.86$). ΔADC_N did not correlate significantly with ΔFTV_4 for all patients in the standard treatment group ($\rho=-0.12$, $p=0.94$), for those with pCR ($\rho=-0.07$, $p=0.88$), or for those without pCR ($\rho=0.04$, $p=0.83$). Similarly, ΔADC_N did not correlate significantly with ΔFTV_4 for all patients in the experimental treatment group ($\tau_b=-0.02$, $p=0.86$), for those with pCR ($\tau_b=0.10$, $p=0.47$), or for those without pCR ($\tau_b=0.01$, $p=0.94$).

Absolute change in ADC (ΔADC_A) for the full cohort did not correlate with ΔFTV_2 ($\rho=0.03$, $p=0.84$), nor was there a correlation in the full cohort when pCR was ($\rho=0.19$,

p=0.27) or was not (ρ=-0.12, p=0.34) observed. ΔADC_A and ΔFTV_2 were not correlated for patients of the standard drug arm (ρ=0.03, p=0.84). A trend for significant correlation was found between ΔADC_A and ΔFTV_2 (ρ=0.69, p=0.07) when considering patients who received the standard treatment with pCR (n=8), where a strong correlation was found for ΔADC_N. This correlation was not significant for patients of the standard drug arm without pCR (ρ=-0.10, p=0.58). ΔADC_A and ΔFTV_2 were not correlated for those that received the experimental drug (ρ=-0.07, p=0.61), or for patients with (ρ=0.14, p=0.48) or without (ρ=-0.16, p=0.35) pCR that received the experimental drug.

The correlation of ΔADC_A with ΔFTV_4 was not significant for the full cohort (τ_b=0.08, p=0.27). For patients from the full cohort that exhibited pCR (n=34), there was a mild but significant correlation of ΔADC_A and ΔFTV_4 (τ_b=0.25, 95% CI [0.06, 0.44], p=0.044), that was not significant when pCR was not observed (τ_b=0.05, p=0.59). ΔADC_A was not significantly correlated with ΔFTV_4 in patients who received the standard therapy (τ_b=0.03, p=0.78), and was not correlated in patients of the standard therapy group with (τ_b=0.07, p=0.90) or without (τ_b=0.02, p=0.86) pCR. For those that received the experimental drug (n=61), there was a trend for slight correlation between ΔADC_A and ΔFTV_4 (τ_b=0.16, p=0.08). Those that received the experimental therapy and exhibited pCR (n=26) also demonstrated a significant correlation (τ_b=0.479, 95% CI [0.10, 0.53], p=0.015) between ΔADC_A and ΔFTV_4 that was not observed without pCR (τ_b=0.11, p=0.38).

Secondary Aim 3: ΔADC_N *in responders and non-responders*

In the full cohort, the Mann-Whitney *U* test showed a significant mean difference in ΔADC_N at 9.74% (95% CI [2.24, 17.513], p=0.012) in responders compared to non-responders (*figure 7*). Though the mean ΔADC_N differences between responders and non-responders in the standard (11.43%, 95% CI [-1.35, 23.38], p=0.10) and experimental (7.88%, 95% CI [-2.70, 18.33], p=0.13) treatment groups were not statistically significant, higher ΔADC_N was observed in the responder groups with p values approaching 0.05.

Using Response$_{FTV}$ as an outcome, no significant differences in ΔADC_N were observed to distinguish responders and non-responders in the full cohort (2.28%, 95% CI [-5.66, 9.87], p=0.60), standard (-2.18%, 95% CI [-17.83, 11.39], p=0.76), and experimental (4.13%, [-5.98, 13.94], p=0.41) treatment groups.

Discussion:

Results of this study showed that ΔADC_N and ΔFTV_2 had similar predictive performance. Though FTV over the course of treatment is one of the current standard metrics for prediction of recurrence free survival, ΔADC_N may add value as a covariate with ΔFTV_2 for prediction of treatment response. Although ΔFTV_4 was not a satisfactory alternative outcome in this population, FTV has been utilized as a surrogate endpoint in larger patient populations and with multiple chemotherapy combinations. Further investigation of this phenomenon is needed.

Although the overall distribution of ΔADC_N in this cohort was not substantially

affected by the quality scoring, better control of ROI delineation, fat suppression, SNR and susceptibility of artifacts would further limit the variation of ADC values in individual patients.   More stringent quality scoring measures may be implemented to further reduce the variation in image quality and to further understand both the limitations of quality scoring via visual examination and the benefits of the methods of ΔADC_N derivation used in this study.

Though ΔADC_N strongly correlated with ΔFTV_4 in patients who received the standard treatment and exhibited pCR, the small sample size (n=8) and wide confidence interval (95% CI [0.22, 1.00]) indicate that this result is likely clinically insignificant. The overall lack of correlation between ΔADC_N and ΔFTV at both early and late time points and presence of correlations between ΔADC_A and ΔFTV_4 that are only significant in populations with pCR indicate that ADC and FTV may account for distinct physiological processes that provide complementary information. Therefore, the diffusivity information obtained from DWI can be combined with volumetric data from DCE-MRI, ultimately increasing the ability to predict response earlier in treatment.

A possible limitation of this study is the use of SS-EPI acquisition methods. EPI techniques used most likely increased the variation of ADC values, as they result in lower spatial resolution, ghosting artifacts, susceptibility artifacts and lower signal intensity due to T2 and T2* decay. Since a minimum of four image series are required for DWI, the effect of respiratory and other physiologic motion can be amplified. The averaging of the signal intensities from each of the gradient directions (*figure 4B, C*), can cause blurring, ultimately diminishing image quality. However, the incorporation of

EPI increases the speed of acquisition and reduces the impact of these artifacts. The use of parallel imaging further advantages in terms of SNR and speed, despite a loss in signal intensity, that allows SS-EPI to be an efficient DWI acquisition method [24]. While the I-SPY 2 imaging protocol is meant to minimize differences in MRI acquisition, it is likely that some of the variation in ADC values can be attributed to EPI techniques.

The predictive ability of ADC measurements or change in ADC has been explored in previous studies. However, small sample sizes [6,12-13,25] and low prevalence of pCR [25] limit their utility. I-SPY 2 is a large multisite trial that implemented strict and structured protocols to provide complete FTV data with standardized treatment and pathologic endpoints. With an increase in sample size an implementation of image quality control, the benefits of DWI can be more accurately assessed.

Since metrics such as categorical quality scoring and the benefits of ADC value derivation based on more than two prescribed b-values have not been fully vetted, the most pressing priority will be to assess the availability of these extra DWI scans and standardize the protocol. The lack of standardization for acquired b-values will affect the noise level of the DWI, and consequently will alter the accuracy of the derived ADC maps. Though determination of the distribution of the b-values in the dataset used in this project may prove useful in the optimization of this value, only 7% of exams acquired b-values other than 0 and 800 s/mm^2.

ACRIN 6698, a sub study of I-SPY 2 aimed at assessing the prediction of response using four b-values in the measurement of ADC values at each MR exam, began enrolling patients shortly after the completion of the MR2 exams from this study in late

August 2012. The ACRIN 6698 sub study also incorporated DWI quality control measures such as protocol compliance monitoring (https://www.acrin.org/6698_protocol.aspx) and quality scoring. The optimal ROI delineation, for malignant and normal tissue alike, is still under investigation, and ACRIN 6698 aims to develop ROI methodologies and determine the optimal b-values for ADC mapping.

Conclusion:

DWI is a non-contrast alternative to DCE for the visualization of treatment response. It is a promising technique that provides complementary tumor information regarding cell density and diffusivity. These findings suggest that ΔADC_N may be more sensitive to the early effects of neoadjuvant treatment than DCE, the standard imaging biomarker that reflects tumor permeability. ΔADC_N was also robust to variability in DWI quality, indicating that the derivation of ΔADC_N in this case can be implemented without regards to image quality. While there was a lack of correlation with ΔFTV_2 or ΔFTV_4, ΔADC_A correlated with ΔFTV_4 for responders in particular. These results suggest that further study of the clinical benefits of DWI and the standardization of b-value prescription is warranted. Areas for further study include investigation based on subtypes, the study of image quality scoring benefits, comparison of normal and tumor ADC values in the ipsilateral breast alone, and establishing objective, automated techniques for obtaining tumor ADC values for clinical purposes.

References:

1. Hylton NM, *et al.* "Locally advanced breast cancer: MR imaging for prediction of response to neoadjuvant chemotherapy—results from ACRIN 6657/I-SPY TRIAL." *Radiology,* 263.3 (2012): 663-672.

2. Woodhams R, Matsunaga K, Kan S, Hata H, Ozaki M, Iwabuchi K, Kuranami M, Watanabe M, Hayakawa K. "ADC Mapping of benign and malignant breast tumors." *Magn. Reson. Med. Sci*. 4 (2005): 35–42.

3. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ. "I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy." *Clin Pharmacol Ther*. 86.1 (2009): 97-100.

4. Solomon L, *et al.* "Veliparib (ABT-888) potentiates the cytotoxic activity of DNA alkylating agents by trapping PARP onto damaged chromatin." *Cancer Research* 74.19 Supplement (2014): 2734-2734.

5. Slamon DJ, *et al.* "Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene." *Science.* 235.4785 (1987): 177-182.

6. Imamura T, *et al.* "Diagnostic performance of ADC for Non-mass-like breast lesions on MR imaging." *Magnetic Resonance in Medical Sciences* 9.4 (2010): 217-225.

7. Tamura T, *et al.* "Investigation of the optimal b value to detect breast tumors with diffusion weighted imaging by 1.5-T MRI." *Cancer Imaging* 14.1 (2014): 11.

8. Partridge SC, DeMartini WB, Kurland BF, Eby PR, White SW, Lehman CD. "Quantitative diffusion-weighted imaging as an adjunct to conventional breast

MRI for improved positive predictive value." *Am. J. Radiology*. 193 (2009):1716–1722.

9. 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).

10. Mukherjee P, *et al.* "Diffusion tensor MR imaging and fiber tractography: theoretic underpinnings." *American journal of neuroradiology,* 29.4 (2008): 632-641.

11. Partridge SC, McDonald ES. "Diffusion Weighted MRI of the Breast: Protocol Optimization, Guidelines for Interpretation, and Potential Clinical Applications." *Magnetic resonance imaging clinics of North America.* 21.3 (2013): 601–624.

12. Pickles M, Gibbs P, Lowry M, Turnbull L. "Diffusion changes precede size reduction in neoadjuvant treatment of breast cancer." *Magn. Reson. Imaging.* 24 (2006): 843–847.

13. Sharma U, Danishad KK, Seenu V, Jagannathan NR. "Longitudinal study of the assessment by MRI and diffusion-weighted imaging of tumor response in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy." *NMR Biomed.* 22 (2009):104–113.

14. Yankeelov TE, Lepage M, Chakravarthy A, Broome EE, Niermann KJ, Kelley MC, Meszoely I, Mayer IA, Herman CR, McManus K, Price RR, Gore JC. "Integration of quantitative DCE-MRI and ADC mapping to monitor treatment response in human breast cancer: initial results." *Magn. Reson. Imaging.* 25 (2007): 1–13.

15. EI Khouli RH. *et al.* "Diffusion-Weighted Imaging Improves the Diagnostic Accuracy of Conventional 3.0-T Breast MR Imaging." *Radiology,* 256.1 (2010): 64–73. *PMC.*

16. Partridge SC, Gibbs JE, Lu Y, Esserman LJ, Tripathy D, Wolverton DS, Rugo HS, Hwang ES, Ewing CA, Hylton NM. MRI measurements of breast tumor volume predict response to neo-adjuvant chemotherapy and recurrence-free survival. AJR Am. J. Roentgenol. 184 (2005): 1774–1781.

17. Hylton NM. "Vascularity assessment of breast lesions with gadolinium-enhanced MR imaging." *Magn. Reson. Imaging Clin. N. Am*, 7.2 (1999): 411–420, x.

18. Newitt DC, *et al.* "Real-Time Measurement of Functional Tumor Volume by MRI to Assess Treatment Response in Breast Cancer Neoadjuvant Clinical Trials: Validation of the Aegis SER Software Platform." *Translational Oncology*. 7.1 (2014): 94–100.

19. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Mller M. "pROC: an open-source package for R and S+ to analyze and compare ROC curves." *BMC Bioinformatics*. 12 (2011): 77.

20. DeLong ER, *et al.* "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach." *Biometrics,* (1988): 837-845.

21. Cardoso F, Van't Veer L, Rutgers E, Loi S, Mook S, Piccart-Gebhart MJ. "Clinical application of the 70-gene profile: the MINDACT trial." *J. Clin. Oncol*, 26. (2008): 729–735.

22. Mook S, Van't Veer LJ, Rutgers, EJ, Piccart-Gebhart MJ, Cardoso F. "Individualization of therapy using MammaPrint: from development to the MINDACT Trial." *Cancer Genomics Proteomics,* 4 (2007): 147–155.

23. Wolf, D. M., et al. "Abstract P1-08-01: MammaPrint ultra-high risk score is associated with response to neoadjuvant chemotherapy in the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657)." *Cancer Research* 73.24 Supplement (2013): P1-08.

24. Mukherjee P, *et al*. "Diffusion tensor MR imaging and fiber tractography: technical considerations." *American Journal of Neuroradiology,* 29.5 (2008): 843-852.

25. Wu L, Chang R, Huang C, Lu Y, Chen H, Chen J, Chang Y. "Evaluation of the treatment response to neoadjuvant chemotherapy in locally advanced breast cancer using combined magnetic resonance vascular maps and apparent diffusion coefficient." *J. Magn. Reson. Imaging.* (2015) 1522-2586.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature _____   Date   8/26/15_____