Diffusion-Weighted MRI for Predicting Pathologic Complete Response in Neoadjuvant Immunotherapy

Wen Li,†, Nu N. Le, Natsuko Onishi, David C. Newitt, Lisa J. Wilmes, Jessica E. Gibbs, Julia Carmona-Bozo, Jiachao Liang, Savannah C. Partridge, Elissa R. Price, Bonnie N. Joe, John Kornak, Mark Jesus M. Magbanua, Rita Nanda, Barbara LeStage, Laura J. Esserman, I-SPY Imaging Working Group, I-SPY Investigator Network, Laura J. van’t Veer, and Nola M. Hylton

Abstract: This study tested the hypothesis that a change in the apparent diffusion coefficient (ADC) measured in diffusion-weighted MRI (DWI) is an independent imaging marker, and ADC performs better than functional tumor volume (FTV) for assessing treatment response in patients with locally advanced breast cancer receiving neoadjuvant immunotherapy. A total of 249 patients were randomized to standard neoadjuvant chemotherapy with pembrolizumab (pembro) or without pembrolizumab (control). DCE-MRI and DWI, performed prior to and 3 weeks after the start of treatment, were analyzed. Percent changes of tumor ADC metrics (mean, 5th to 95th percentiles of ADC histogram) and FTV were evaluated for the prediction of pathologic complete response (pCR) using a logistic regression model. The area under the ROC curve (AUC) estimated for the percent change in mean ADC was higher in the pembro cohort (0.73, 95% confidence interval [CI]: 0.52 to 0.93) than in the control cohort (0.63, 95% CI: 0.43 to 0.83). In the control cohort, the percent change of the 95th percentile ADC achieved the highest AUC, 0.69 (95% CI: 0.52 to 0.85). In the pembro cohort, the percent change of the 25th percentile ADC achieved the highest AUC, 0.75 (95% CI: 0.55 to 0.95). AUCs estimated for percent change of FTV were 0.61 (95% CI: 0.39 to 0.83) and 0.66 (95% CI: 0.47 to 0.85) for the pembro and control cohorts, respectively. Tumor ADC may perform better than FTV to predict pCR at an early treatment time-point during neoadjuvant immunotherapy.

Keywords: diffusion-weighted MRI; breast cancer; immunotherapy; pathologic complete response; neoadjuvant therapy
1. Introduction

The number of clinical trials that investigate immunotherapeutic agents for breast cancer treatment has increased in recent years [1–3]. A notable result published from the I-SPY TRIAL (Investigation of Serial studies to Predict Your Therapeutic Response with Imaging And moLecular Analysis) showed that the addition of pembrolizumab to standard neoadjuvant chemotherapy (NAC) achieved a higher estimated pathologic complete response (pCR) rate than standard NAC alone [3]. Recently, the FDA approved pembrolizumab in combination with chemotherapy for high-risk early-stage triple-negative breast cancer [4].

Immunotherapy works by boosting the patient’s own immune system to fight cancer cells, a mechanism of action which differs from that of conventional NAC. Pembrolizumab is a PD-1 (programmed death-1) inhibitor that triggers the activation of the T-cell immune response against cancer cells [5]. In the neoadjuvant setting, the reported pCR rates for immunotherapy (alone or in combination with standard NAC) range from 43.5% to 64.8% [6], which means approximately half of the patients treated with immunotherapy do not achieve pCR. Studies showed that the patients who achieve pCR after NAC had better recurrence-free survival than patients who do not achieve pCR [7–9]. Since not all of the patients benefit from immunotherapy and some patients may experience serious side effects [6,10], it is crucial to assess tumor response early during treatment. However, assessing the response to immunotherapy using conventional imaging methods, such as measuring changes in tumor size according to the Response Evaluation Criteria in Solid Tumors (RECIST), is limited because RECIST was designed to assess the response to standard therapies [11,12]. Pseudo-progression with immunotherapy is often observed, and is characterized by an initial increase in tumor size followed by a decrease in tumor size or volume [13–15].

Diffusion-weighted MRI (DWI) is a non-invasive, non-contrast MRI technique that is different from dynamic contrast-enhanced (DCE) MRI. The apparent diffusion coefficient (ADC) calculated from DWI, reflects the water diffusion within the tumor microenvironment and the ADC values are associated with the cellularity and integrity of cell membranes. Results from the multi-center ACRIN 6698 clinical trial showed that the percent change of breast tumor ADC measured 12 weeks after the start of NAC is predictive of pCR [16]. A subsequent multi-center study showed that ADC adds independent value to functional tumor volume (FTV) measured from DCE-MRI in the prediction of pCR [17]. FTV is a well-established imaging marker of treatment response demonstrated by the multi-center I-SPY 1/ACRIN 6657 clinical trial [18,19]. In I-SPY 2, FTV continues to be used as an integral imaging marker for monitoring treatment response. Each patient who enrolls in the I-SPY 2 trial undergoes MRI scans including both DCE and DWI at four treatment time-points. However, the optimum use of DCE and/or DWI in assessing response to immunotherapy is still unclear. In this study, we compared the predictive performance of ADC with that of FTV, both measured at an early treatment time-point, for patients treated with pembrolizumab in addition to standard chemotherapy and for patients treated with standard chemotherapy alone in I-SPY 2. As a secondary aim, we explored the histogram percentiles as an alternative ADC metric other than mean ADC to address tumor heterogeneity as a potential imaging marker to predict pCR.

2. Materials and Methods

2.1. Patient Population

The patient population of this study was a sub-cohort from the I-SPY 2 trial, an adaptively randomized phase 2 neoadjuvant clinical trial for early-stage high-risk breast cancer (ClinicalTrials.gov ID: NCT01042379) [20]. I-SPY 2 is a multi-site clinical trial and each participating institute obtained Institutional Review Board (IRB) approval independently. Written consents for both the screening and treatment were required for each patient. Patients with HER2-negative breast cancer who enrolled in I-SPY 2 consecutively between 26 November 2015 and 5 November 2016 were eligible to be randomized to paclitaxel...
+ pembrolizumab for 12 weeks followed by four cycles of doxorubicin plus 600 mg/m² intravenous cyclophosphamide (AC) every 2 to 3 weeks ($n = 69$, referred to as the “pembro” arm below). The control group consisted of 180 patients who had HER2-negative breast cancer and were randomized to receive standard chemotherapy (paclitaxel followed by AC). The sample sizes of the control and pembro groups were determined by the I-SPY 2 drug graduation rule. Details of the study design related to this patient population in I-SPY 2 were published previously [21].

2.2. MRI Analysis

Breast MRIs were performed at participating institutions on whole body MRI scanners from different vendors, at field strengths of 1.5 T or 3.0 T using a dedicated breast coil. Studies at baseline/prior to treatment (T0) and 3 weeks after start of treatment (T1) were used for this analysis of early prediction of treatment response. Additional MRI exams performed at inter-regimen (approximately 12 weeks after the therapy started) and before surgery (approximately 24 weeks after the therapy started) were not included in this analysis. The same scanner, MRI breast coil, and sequences were used to scan the same patient across the multiple time-points. Details of the I-SPY 2 MRI protocol were previously published [17]. Briefly, the I-SPY 2 MRI scan protocol used for this cohort included bilateral axial acquisition of a T2-weighted sequence, a pre-contrast echo planar imaging (EPI) DWI sequence, and DCE using a bilateral, 3D, fat-suppressed, and T1-weighted sequence with 80–100 s temporal resolution. DWI was acquired using two $b$-values: 0 and 800 s/mm² with an acquisition time ≤ 5 min.

ADC maps were calculated from DWI using in-house software developed in IDL (Exelis Visual Information Solutions, Boulder, CO, USA) based on the monoexponential decay model: $ADC = -\ln \frac{S_0}{S} / b$, where $b = 800$ s/mm² and $S_0$ is the signal intensity corresponding to no diffusion gradients ($b = 0$ s/mm²). The DWI data and corresponding ADC maps were reviewed, and the tumor region-of-interest (ROI) was manually drawn by two readers with 7 (W.L.) and 0.2 (N.L.) years of experience in breast DWI analysis. Both readers were blinded to the pathologic outcomes. Each reader was trained on a separate cohort of 30 patients with DWI acquired at baseline and 3-week time-points. Discrepancies between the readers were resolved after consultation of a breast radiologist with 10 years of experience (N.O.). After the training, both of the readers assessed all of the DWI in this study independently and the corresponding report of the reader study was published [22]. The ROIs in the present study were multiple-slice restricted ROIs, defined as the most diffusion-restricted area on DWI (dark on ADC map and bright on $b = 800$ s/mm² diffusion-weighted image) in all of the axial slices in which the tumor can be identified [22]. Based on the agreement found in the reader study, multiple-slice restricted ROIs from one reader (W.L.) were used in this study. To avoid including volumes with T2 shine-through effect, the ROI placement was strictly limited to the volume with high signal intensity on $b = 800$ s/mm² diffusion-weighed image with low ADC values. In addition, special care was taken to avoid cystic, necrotic or fatty components by excluding areas with no contrast enhancement on DCE-MRI [23]. The ROIs at 3-week were placed at similar locations to the baseline MRI, if possible, given the differences in breast position and treatment response, to be able to reflect the treatment response. ADC metrics were quantified using estimated mean and percentiles (minimum, 5th, 15th, 25th, 50th, 75th, 95th, maximum) from ADC histograms extracted from ROIs. In addition, DWIs were ranked based on image quality in both $b = 0$ and $b = 800$ images based on fat suppression, artifacts, and signal-to-noise ratio.

FTV (in cubic centimeter, cc) was also calculated using DCE-MRI as a standard imaging marker in I-SPY 2. Specifically, percent enhancement (PE, in %) and signal enhancement ratio (SER) maps were derived from DCE-MRI using T1-weighted images acquired before and multiple times after the injection of contrast agent. PE was calculated as the percent change of signal intensity between pre-contrast and early (nominal 2.5 min after injection)
post-contrast: \( PE = \frac{S_1 - S_0}{S_0} \times 100 \). SER was calculated as the ratio of early enhancement to late (nominal 7.5 min after) enhancement: \( SER = \frac{S_1 - S_0}{S_2 - S_0} \), where \( S_0 \), \( S_1 \), and \( S_2 \) represent the signal intensities of each voxel in the pre-contrast, early post-contrast, and late post-contrast images, respectively. FTV was computed as the combined volume of all voxels with \( PE \geq 70\% \) and \( SER \geq 0 \).

### 2.3. Treatment Outcome

After completion of NAC, patients underwent surgery (lumpectomy or mastectomy). The surgical specimens were analyzed for response by site pathologists trained to assess residual invasive disease, with pCR defined as no invasive residual disease in primary breast and lymph nodes. The binary pathological outcome (pCR versus non-pCR) was used as the treatment outcome in this study. The patients who exited the trial without surgery or did not undergo surgery for various reasons were excluded from our analysis.

### 2.4. Statistical Analysis

Multiple MRI tumor metrics, calculated from DWI and DCE-MRI, were evaluated for the prediction of pCR. These metrics were: percent change of ADC metrics (mean, minimum, maximum, as well as 5th, 15th, 25th, 50th, 75th, 95th percentile) and percent change of FTV between T0 and T1 were considered as potential MRI variables (herein referred to as ADC_mean, ADC_min, ADC_max, ADC_5, ADC_15, ADC_25, ADC_50, ADC_75, ADC_95, FTV, respectively) for the prediction of pCR. To enable a direct comparison between ADC and FTV, the percent change of ADC was calculated as: \([ \text{ADC at T1} – \text{ADC at T0}] / \text{ADC at T0} \). The percent change of FTV was calculated as: \([ \text{FTV at T1} – \text{FTV at T0}] / \text{FTV at T0} \). According to our observation, the distribution of percent change of ADC and percent change of FTV were skewed so their values were given as a median with the interquartile range. The associations between MRI variables and pCR were analyzed by single predictor or multiple predictors in a logistic regression model. The area under the receiver operating characteristic curve (AUC) was used to assess the predictive performance of a single-predictor model. Hypothesis tests of the differences between two AUCs were evaluated using the DeLong test [24]. The \( p \)-values of the differences between MRI variables in pCR versus non-pCR groups were estimated using the Wilcoxon rank sum test. The impact of FTV and treatment on the association between ADC with pCR was studied using nested models where FTV and treatment were added one-at-a-time as covariate to the model starting with ADC alone. The odds ratio of percent change of ADC or FTV for pCR was calculated per 10% increments in the logistic regression model. Correlation coefficients and their significant levels (\( p \)-values) between two MRI variables were tested by Spearman’s rank correlation test. To compare the predictive performance of MRI variables, the full analysis cohort was stratified by treatment arm: control (patients received paclitaxel only as the first regimen) and pembro (patients received paclitaxel + pembrolizumab as the first regimen). Statistical analysis was implemented in R version 4.1.3 and the statistical significance level was set at \( \alpha = 0.05 \).

### 3. Results

A total of 249 patients were considered for inclusion in the analysis. One hundred and eighty patients were randomized to be the controls and 69 patients to the pembro. Figure 1 shows a flowchart of data exclusion and inclusion. In summary, 103 total patients were included in the analysis (75 in controls and 28 in pembro) and the majority of the exclusions (117 patients in total) are due to the poor quality of DWI at either or both T0 and T1. The patient characteristics of the excluded and included cohorts are shown in Table 1. The included patients were on average 3 years younger than those who were excluded (\( p = 0.004 \)). Race, HR/HER2 subtype, treatment, or menopausal status between the included and excluded cohorts were balanced between the two groups.
Figure 1. Data inclusion and exclusion. Out of 249 patients who had HER2- breast cancer, 103 patients were included in our analysis.

Table 1. Patient characteristics.

|                      | Included (N = 103) | Excluded (N = 146) | p *  |
|----------------------|-------------------|-------------------|-----|
| Age                  | 46 +/- 11         | 49 +/- 11         | 0.004 |
| Race                 |                   |                   |     |
| White                | 76 (74%)          | 115 (79%)         | 0.54 |
| Black or African American | 15 (15%)      | 17 (12%)          |     |
| Asian                | 8 (8%)            | 8 (6%)            |     |
| American Indian or Alaska Native | 0 (0%)    | 3 (2%)            |     |
| Others               | 3 (3%)            | 3 (2%)            |     |
| Subtype              |                   |                   | 0.12 |
| HR+                  | 62 (60%)          | 72 (49%)          |     |
| HR-                  | 41 (40%)          | 73 (50%)          |     |
| Treatment            |                   |                   | 0.89 |
| Paclitaxel           | 75 (73%)          | 105 (72%)         |     |
| Paclitaxel + Pembrolizumab | 28 (27%)  | 41 (28%)          |     |
| Menopausal Status    |                   |                   | 0.26 |
| Premenopausal        | 58 (56%)          | 68 (47%)          |     |
| Postmenopausal       | 26 (25%)          | 49 (34%)          |     |
| Perimenopausal       | 7 (7%)            | 6 (4%)            |     |
| Others               | 12 (12%)          | 23 (16%)          |     |

* Assessed by Student’s t-test for age and Fisher’s test for categorical variables.

Percent changes of ADC metrics at T1 compared to T0 are shown in Table 2. The values of ADC metrics and FTV at T0 and T1 can be found in the Supplementary Materials, Tables S1 and S2. In general, the percent changes in tumor ADC were positive, which indicates that ADC increased after 3 weeks of treatment. In the full cohort of 103 patients, statistically significant differences between the pCR and non-pCR groups were observed in ADC_mean, ADC_min, and lower percentiles of ADC histogram—up to the 50th percentile. In the control cohort, a statistically significant difference was only observed at ADC_95. In the pembro cohort, statistically significant differences were observed at ADC_mean, ADC_min, ADC_15, and ADC_25. Percent changes of FTV in the full cohort and in individual treatment cohorts are listed in Table 2. FTV decreased in all of the cohorts despite pCR outcomes. The difference between pCR and non-pCR groups reached statistically significance (p = 0.016) in the full cohort but not in the sub-cohorts.

Figure 2 shows boxplots of the percent change of ADC and FTV at T1 by pCR outcome and cohort. Although both the ADC and FTV indicated a good separation between the pCR and non-pCR groups in the full cohort, ADC was visually better separated than FTV in the pembro cohort with p-values of 0.041 and 0.34 for ADC and FTV, respectively (Table 2). The boxplots suggested stronger differences were found for ADC_mean between the pCR and non-pCR groups.
Table 2. Percent change MRI variables after 3 weeks of therapy.

| MRI Variable | Full (n = 103) | Control (n = 75) | Pembro (n = 28) |
|--------------|---------------|-----------------|-----------------|
|              | pCR (n = 30)  | Non-pCR (n = 73) | pCR (n = 15)   | Non-pCR (n = 60) | pCR (n = 15) | Non-pCR (n = 13) | p   |
| ADC_mean     | 18.9 (4.4, 23.6) | 10.2 (2.6, 15.3)  | 0.0076         | 15.8 (3.1, 22.9) | 10.3 (5.1, 15.8) | 0.13         | 19.2 (8.2, 23.1) | 3.5 (−2.4, 14.7) | 0.041 |
| ADC_min      | 28.0 (5.9, 54.7) | 12.6 (−0.7, 24.6) | 0.0075         | 23.4 (−2.7, 53.0) | 14.1 (−0.8, 19.4) | 0.15         | 29.7 (10.6, 56.5) | 9.6 (−0.2, 20.8) | 0.029 |
| ADC_5        | 20.0 (−1.1, 32.8) | 8.0 (0, 17.5)    | 0.0297         | 18.2 (−6.5, 33.5) | 9.1 (−0.3, 19.4)  | 0.25         | 23.5 (4.7, 32.1)  | 6.1 (4.6, 10.5)  | 0.0799 |
| ADC_15       | 17.3 (0.3, 27.6) | 9.6 (2.4, 16.0)  | 0.039          | 14.6 (0.6, 26.7)  | 9.9 (3.3, 17.2)   | 0.47          | 22.2 (2.2, 28.3)  | 7.8 (1.1, 11.2) | 0.040  |
| ADC_25       | 17.9 (3.1, 29.7) | 8.6 (2.4, 15.4)  | 0.014          | 15.7 (1.8, 23.2)  | 10.2 (3.6, 16.2)  | 0.25          | 23.6 (5.3, 31.4)  | 5.2 (0, 11.9)    | 0.025  |
| ADC_50       | 18.0 (3.0, 23.5) | 10.1 (3.2, 15.5) | 0.048          | 13.2 (0.2, 23.8)  | 10.4 (3.8, 15.5)  | 0.34          | 19.5 (7.1, 22.96) | 2.9 (−1.2, 14.2) | 0.098  |
| ADC_25       | 14.9 (1.5, 23.7) | 8.5 (1.8, 15.6)  | 0.052          | 14.6 (1.1, 25.3)  | 8.9 (2.8, 14.9)   | 0.19          | 15.2 (5.5, 21.1)  | 1.8 (−3.6, 17.4) | 0.196  |
| ADC_95       | 15.0 (1.1, 20.4) | 7.3 (0, 14.0)    | 0.087          | 18.6 (7.9, 22.6)  | 7.5 (1.5, 14.3)   | 0.027         | 6.01 (0.5, 16.1)  | 0.8 (−4.98, 12.5) | 0.39   |
| ADC_max      | 6.3 (−1.2, 11.8) | 7.4 (−1.7, 13.8) | 0.85           | 10.8 (1.8, 14.8)  | 7.5 (−2.2, 14.6)  | 0.53          | 3.8 (−5.5, 9.7)   | 3.8 (0.7, 9.9)   | 0.82   |

FTV = −56.97 (−83.9, −19.98) −31.5 (−58.4, −3.8) 0.016 −59.9 (−82.0, −14.2) −30.6 (−54.2, −2.7) 0.056 −47.7 (−83.4, −20) −46.3 (−64.6, −8.99) 0.34

Figure 2. Boxplots of percent change in mean ADC and FTV at 3-week. The percent change in FTV was multiplied by −1 so a positive change in both ADC and FTV means positive response to treatment. Each dot represents an individual patient. ADC: apparent diffusion coefficient. FTV: functional tumor volume. pCR = 0 are non-pCRs and pCR = 1 are pCRs.
The AUCs for the percent change of the ADC metrics evaluated to predict pCR are listed in Table 3. In the full cohort, the AUCs ranged from 0.49 to 0.67, with the highest AUCs for ADC_mean and ADC_min. When the full cohort was stratified by treatment, the AUCs ranged from 0.55 to 0.69 and 0.47 to 0.75 for the control and pembro cohorts, respectively. The highest AUCs were observed for ADC_95 and ADC_25 percentiles in the control and pembro cohorts, respectively. Although the highest AUCs were not observed at ADC_mean, the difference between the ADC_mean and ADC metrics that yielded the highest AUCs did not reach the pre-defined statistical significance level either in the control (AUC = 0.63 versus AUC = 0.69, \( p = 0.43 \)) cohort or in the pembro (AUC = 0.73 versus AUC = 0.75, \( p = 0.62 \)) cohort.

Table 3. Area under the ROC curve for percent change of MRI variables at 3-week.

| MRI Variable | Full (n = 103, pCR Rate 29%) | Control (n = 75, pCR Rate 20%) | Pembro (n = 28, pCR Rate 54%) |
|--------------|------------------------------|--------------------------------|-----------------------------|
| ADC_mean     | 0.67 (0.53, 0.81)            | 0.63 (0.43, 0.83)              | 0.73 (0.52, 0.93)           |
| ADC_min      | 0.67 (0.53, 0.80)            | 0.62 (0.42, 0.82)              | 0.74 (0.55, 0.93)           |
| ADC_5        | 0.64 (0.49, 0.78)            | 0.58 (0.38, 0.78)              | 0.697 (0.48, 0.91)          |
| ADC_15       | 0.63 (0.49, 0.77)            | 0.56 (0.36, 0.76)              | 0.73 (0.52, 0.94)           |
| ADC_25       | 0.66 (0.51, 0.80)            | 0.597 (0.397, 0.797)           | 0.75 (0.55, 0.95)           |
| ADC_50       | 0.62 (0.49, 0.76)            | 0.58 (0.38, 0.78)              | 0.69 (0.48, 0.897)          |
| ADC_75       | 0.62 (0.49, 0.76)            | 0.61 (0.41, 0.81)              | 0.65 (0.43, 0.86)           |
| ADC_95       | 0.61 (0.48, 0.74)            | 0.69 (0.52, 0.85)              | 0.60 (0.38, 0.82)           |
| ADC_max      | 0.49 (0.37, 0.61)            | 0.55 (0.39, 0.71)              | 0.47 (0.24, 0.70)           |
| FTV          | 0.65 (0.52, 0.78)            | 0.66 (0.47, 0.85)              | 0.61 (0.39, 0.83)           |

The AUCs for the percent change of FTV were estimated as 0.65 (95% CI: 0.52 to 0.78), 0.66 (95% CI: 0.47 to 0.85), and 0.61 (95% CI: 0.39 to 0.83) in the full cohort, control, and pembro, respectively (Table 3). Figure 3 displays the ROC curves for ADC and FTV. The biggest separation of two ROC curves was observed in the pembro cohort, suggesting a higher performance for ADC in this cohort.

Figure 3. ROC curves of percent change in mean ADC and FTV for predicting pCR at 3-week. In the full cohort (left), pCR rate is 29% (n = 30) and the area under the ROC curve is 0.67 for ADC and 0.65 for FTV. In the control cohort (middle), pCR rate is 20% (n = 15) and the area under the ROC curve is 0.63 for ADC and 0.66 for FTV. In the pembro cohort (right), pCR rate is 54% (n = 15) and the area under the ROC curve is 0.73 for ADC and 0.61 for FTV. ADC: apparent diffusion coefficient. FTV: functional tumor volume. pCR: pathologic complete response.

Table 4 shows the odds ratios, 95% CI thereof, and the \( p \)-values for ADC_mean in the logistic regression model of predicting pCR, where ADC_mean was a single predictor (left column), FTV was added (middle column), and FTV and treatment were both added.
to the model (right column). The odds ratios were estimated in the full cohort \((n = 103, \text{pCR rate} = 29\%)\). Table 4 shows that ADC_mean had a statistically significant association with pCR even in the multivariate models. For comparison, the effect of ADC_mean and treatment to FTV is shown in the last row of Table 4, where the significance of FTV changed \((p\text{-value from 0.039 to 0.045})\) after ADC_mean was added to the model.

Table 4. Odds ratio of percent change in mean ADC in logistic regression model (outcome: pCR).

| MRI Variable | Univariate | Multivariate: ADC + FTV | Multivariate: ADC + FTV + Treatment |
|--------------|------------|-------------------------|------------------------------------|
|              | Odds Ratio | p                       | Odds Ratio                         | Odds Ratio | p                  |
| Mean ADC     | 1.67 (1.17, 2.5) | 0.0039                 | 1.70 (1.18, 2.59) | 0.0032     | 1.71 (1.18, 2.62) | 0.0032 |
| FTV          | 0.90 (0.80, 0.995) | 0.039                  | 0.90 (0.80, 0.998) | 0.045      | 0.92 (0.81, 1.02) | 0.045  |

The correlations between ADC_mean and FTV were weak and not statistically significant: \(r = 0.041 (p = 0.68)\) in the full cohort, \(r = -0.089 (p = 0.45)\) in control, and \(r = 0.29 (p = 0.13)\) in pembro.

Figures 4 and 5 show example cases of pCR and non-pCR. In the pCR case, the tumor FTV was 163.9cc at T0 and 155.6cc at T1 (5% decrease). ADC_mean was \(1.102 \times 10^{-3} \text{ mm}^2/\text{s}\) at T0 and \(1.437 \times 10^{-3} \text{ mm}^2/\text{s}\) at T1 (30% increase). In the non-pCR case, the tumor FTV was 3.5cc at T0 and 0.6cc at T1 (82% decrease), while mean ADC was \(0.819 \times 10^{-3} \text{ mm}^2/\text{s}\) and \(0.831 \times 10^{-3} \text{ mm}^2/\text{s}\), respectively (1% increase). ADC changed more than FTV in the pCR case and less than FTV in the non-pCR case. Additional example cases from the control cohort are provided in the Supplementary Materials, Figures S1 and S2.

**Figure 4.** Example case 1 with a pathologic complete response. The patient was treated by paclitaxel + pembrolizumab followed by cyclophosphamide. Representative MR images are shown from pretreatment (top row) and at 3-week (bottom row). Images from DCE-MRI were acquired 144 s after the contrast injection. Images from DWI are shown in a pair of original DWI \(b = 800 \text{ s/mm}^2\) and ADC map. ROIs are shown in yellow (rectangular box in DCE and hand-drawn in DWI).
Figure 5. Example case 2 without a pathologic complete response. The patient was treated by paclitaxel + pembrolizumab followed by cyclophosphamide. Representative MR images are shown from pretreatment (top row) and at 3-week (bottom row). Images from DCE-MRI were acquired 119 s after the contrast injection. Images from DWI are shown in a pair of original DWI (b = 800 s/mm$^2$) and ADC map. ROIs are shown in yellow (rectangular box in DCE and hand-drawn in DWI). Images from pretreatment and at 3-week may seem different by visual comparison due to different positioning in the MRI scanner. Breast tissue in DCE-MRI suggests that they are from similar slice location.

4. Discussion

This study measured the percent change of ADC metrics (ADC_mean, ADC_min, ADC_5, ADC_15, ADC_25, ADC_50, ADC_75, ADC_95, maximum) before and after 3 weeks of neoadjuvant therapy in breast cancer patients enrolled in the I-SPY 2 trial. We assessed the prediction of pCR using ADC metrics measured in patients treated with standard NAC alone or in combination with immunotherapy. We also compared the predictive performance of mean ADC with FTV measured from DCE-MRI. We found that all of the ADC metrics except ADC_95 and ADC_max performed better than FTV in the immunotherapy cohort after 3 weeks of therapy.

This study found that the change in the ADC_mean was predictive of pCR. This is consistent with our prior work, in which the ADC_mean was also predictive of pCR at T1 [17]. The previous study also included data from a larger cohort of patients in the I-SPY 2 TRIAL, (n = 328 at T1) but there was no overlap between that cohort and the cohort used in this analysis. The AUC reported previously was 0.57, while it was 0.67 in this analysis. Using a different ROI delineation strategy may have contributed to improved AUC in this study. The use of a single drug arm with controls in this analysis may also have contributed to the improved AUC, since the previous study included patients from four different experimental drug arms in I-SPY 2, with no controls.

In addition to the ADC_mean, we found that changes of lower percentiles (ADC_min, ADC_5, ADC_15, ADC_25, ADC_50) were statistically significant and estimated AUCs were higher for the lower percentiles compared to the higher percentiles. This finding aligns with results reported previously, where lower percentiles seemed more predictive than higher percentiles for treatment response [25,26]. This could be explained by the hypothesis that low ADC may reflect tumor regions of higher cell density that are more sensitive to chemotherapy [27–30]. However, we did not observe this trend in the control cohort. This may indicate response heterogeneity within the tumor.

In the pembro cohort, FTV performed poorly and the majority of the ADC metrics achieved a higher AUC than FTV. However, ADC did not achieve a higher AUC in the control cohort. This finding may indicate that FTV may be limited in assessing the tumor
response to immunotherapy at the early treatment time-point. Previous studies reported that size measurements may not accurately assess the tumor response to immunotherapy due to possible pseudo-progression at early treatment time-points [13,31,32]. Because pseudo-progression has been histologically proved to be infiltration and recruitment of various immune cells including lymphocytes in the tumor, the change of tumor ADC at 3 weeks might reflect the change in cellularity due to the immune cell infiltration, especially the change in lower ADC percentiles (i.e., highly cellular portion).

To our knowledge, this is the first multi-center study to assess tumor ADC in the prediction of treatment response to immunotherapy compared to standard NAC in breast cancer. The results of our study suggest that DWI can be used as a complementary method to DCE-MRI in assessing the response to NAC. However, this study has several limitations. First, the sample size is small, especially in the pembro cohort. The number of patients included in the analysis was 28 (15 pCRs; 13 non-pCRs), which prohibited further analysis by breast cancer subtype. Second, a large number of DWI exams were excluded due to poor image quality, further limiting the sample size for analysis. A total of 117 patients (47%) were excluded because of poor image quality at T0, T1, or both. Most of the excluded cases (n = 102; 32 pembro and 70 controls) had poor-quality DWI in both visits. This high exclusion rate clearly calls for a standardized quality control for DWI. Third, the patients randomized to pembrolizumab were treated by standard NAC plus pembrolizumab so the comparison between the pembro and control cohorts was not purely immunotherapy versus standard NAC. Finally, manual delineation of ROIs and qualitative ranking of DWI quality were subjective and user-dependent. An intelligent and objective quality ranking method for multi-center DWI studies is needed.

5. Conclusions

In conclusion, our findings suggest that at an early treatment time-point, mean ADC performs better than FTV in the prediction of pCR for patients who received neoadjuvant immunotherapy. In addition, we also found that ADC histogram percentiles may achieve a better predictive performance than mean ADC if they were selected based on treatment. Future studies are needed to determine the optimal way of incorporating ADC into the clinical workflow in combination with FTV or replacing FTV as an alternative imaging marker to maximize the prediction of treatment response using MRI in patients receiving immunotherapy.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers14184436/s1, Table S1: tumor ADC metrics and FTV values at T0; Table S2: tumor ADC metrics and FTV values at T1; Figure S1: Example case 3 with a pathologic complete response; Figure S2: Example case 4 without a pathologic complete response.

Author Contributions: Conceptualization, W.L., L.J.W., S.C.P., M.J.M.M., B.L., L.J.E., L.J.v.V. and N.M.H.; methodology, W.L., N.N.L., N.O., S.C.P., E.R.P., B.N.J., J.K. and N.M.H.; software, D.C.N.; formal analysis, W.L.; investigation, W.L., N.N.L., N.O. and R.N.; data curation, J.E.G., E.R.P., B.N.J. and R.N.; writing—original draft preparation, W.L.; writing—review and editing, N.N.L., N.O., D.C.N., L.J.W., J.E.G., J.C.-B., J.L., S.C.P., J.K. and M.J.M.M.; supervision, R.N., L.J.E., L.J.v.V. and N.M.H.; project administration, J.E.G. and L.J.E.; funding acquisition, W.L., M.J.M.M., L.J.E., L.J.v.V. and N.M.H. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National Cancer Institute of the National Institutes of Health (R01 CA132870, U01 CA225427, P01 CA210961, R01 CA255442, R01 CA190299); and the UCSF Breast Oncology Program Research Development Program.

Institutional Review Board Statement: This study used data from the I-SPY 2 TRIAL. The trial was conducted in accordance with the Declaration of Helsinki. All participating sites received written Review Board approval.

Informed Consent Statement: All patients provided written informed consent for screening and for treatment randomization.
**Data Availability Statement:** Imaging and clinical data from this study will be made available upon request in accordance with I-SPY 2 Data and Publication Committee policies.

**Acknowledgments:** The authors would like to thank all patients who participated in the I-SPY 2 TRIAL. We thank the I-SPY 2 Imaging Working Group: Haydee Ojeda-Fornier, Mohammad Eghtedari, Kathryn W. Zamora, Stefanie Woodward, Heidi Umphrey, Michael Nelson, An Church, Patrick Bolan, Theresa Kuritz-Kaye, Katherine Ward, Kevin Morley, Dulcy Wolverton, Kelly Fountain, Dan Lopez-Paniagua, Lara Hardesty, Kathleen R. Brandt, Elizabeth S. McDonald, Mark Rosen, Despina Kontos, Hiroyuki Abe, Deepa Sheth, Erin Crane, Charlotte Dillis, Pulin Sheth, Linda Hovanessian-Larsen, Dae Hee Bang, Bruce Porter, Karen Y. Oh, Neda Jafarian, Luminita A. Tudorica, Bethany Niell, Jennifer Drukteinis, Marina Giurescu, Elisa Berman, Constance Lehan, Kimberly A. Fitzpatrick, Marisa H. Borders, Wei T. Yang, Basak Dogan, Sally Goudreau, Thomas Chenewert and the I-SPY 2 Investigator Network: Angela M. DeMichele, Christina Yau, Claudine Isaacs, W. Fraser Symmans, A. Jo Chien, John W. Park, Hope S. Rubio, Anne M. Wallace, Richard Schweb, Andres Forero-Torres, Erica Stringer-Reason, Douglas Yee, Kathy S. Albain, Anthony D. Elias, Judy C. Boughey, Minetta C. Liu, Amy S. Clark, Olufunmilayo Olopade, Julie E. Lang, Janice Lu, Erin D. Ellis, Kathleen Kemmer, Hyou S. Han, Donald W. Northfelt, Kirsten K. Edmiston, Rachel Yung, Rebecca K. Viscusi, Debasis Tripathy, Stacy L. Moulder, Barbara Haley, Qamar J. Khan, Donald A. Berry.

**Conflicts of Interest:** W.L., N.N.L., N.O., L.J.W., J.E.G., J.C., J.L., E.R.P., J.K., M.J.M.M.: declares no conflict of interest; D.C.N.: Kheiron; S.C.P.: Guerbet, GE, Philips; B.N.J.: Kheiron, UpToDate; R.N.: Astrazeneca, BeyondSprint, Fujifilm, Gilead, Infinity, iTeos, Merck, OBI, Oncosec, Seagen, Arvinas, AstraZeneca, Celgene, Corcept Therapeutics, Genentech/Roche, Gilead/Immunomedics, Merck, OBI Pharma, OncoSec, Pfizer, Relay, Seattle Genetics, Sun Pharma, Taiho; B.L.: Abbott, AbbVie, Johnson & Johnson, Teleflex; L.J.E.: Blue Cross/Blue Shield, Merck; L.V.V.: Agenda; N.M.H.: GE, Kheiron.

**References**

1. Esteva, F.J.; Hubbard-Lucey, V.M.; Tang, J.; Pusztai, L. Immunootherapy and Targeted Therapy Combinations in Metastatic Breast Cancer. Lancet Oncol. 2019, 20, e175–e186. [CrossRef]
2. Schmid, P.; Cortes, J.; Pusztai, L.; McArthur, H.; Kümmerle, S.; Bergh, J.; Kümmel, S.; Park, Y.H.; Hui, R.; Harbeck, N.; et al. Pembrolizumab for Early Triple-Negative Breast Cancer. N. Engl. J. Med. 2020, 382, 810–821. [CrossRef] [PubMed]
3. Nanda, R.; Liu, M.C.; Yau, C.; Asare, S.; Hylton, N.; Veer, L.V.; Perlmutter, J.; Wallace, A.M.; Chien, A.J.; Forero-Torres, A.; et al. Pembrolizumab plus Standard Neoadjuvant Therapy for High-Risk Breast Cancer (BC): Results from I-SPY 2. J. Clin. Oncol. 2017, 35, 506. [CrossRef]
4. FDA. FDA Approves Pembrolizumab for High-Risk Early-Stage Triple-Negative Breast Cancer | FDA. Available online: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-pembrolizumab-high-risk-early-stage-triple-negative-breast-cancer (accessed on 7 August 2022).
5. Sunshine, J.; Taube, J.M. PD-1/PD-L1 Inhibitors. Curr. Opin. Pharmacol. 2015, 23, 32–38. [CrossRef] [PubMed]
6. Fountzila, E.; Ignatiadis, M. Neoadjuvant Immunotherapy in Breast Cancer: A Paradigm Shift? Eancermedicalscience 2020, 14, 1147. [CrossRef]
7. Yee, D.; DeMichele, A.M.; Yau, C.; Isaacs, C.; Symmans, W.F.; Albain, K.S.; Chen, Y.-Y.; Krings, G.; Wei, S.; Harada, S.; et al. Association of Event-Free and Distant Recurrence-Free Survival with Individual-Level Pathologic Complete Response in Neoadjuvant Treatment of Stages 2 and 3 Breast Cancer: Three-Year Follow-up Analysis for the I-SPY2 Adaptively Randomized Clinical Trial. JAMA Oncol. 2020, 6, 1–9. [CrossRef]
8. Cortazar, P.; Zhang, L.; Untch, M.; Mehta, K.; Costantino, J.P.; Wolmark, N.; Bonnefoi, H.; Cameron, D.; Gianni, L.; Valagussa, P.; et al. Pathological Complete Response and Long-Term Clinical Benefit in Breast Cancer: The CTNeoBC Pooled Analysis. Lancet 2014, 384, 164–172. [CrossRef]
9. Spring, L.M.; Fell, G.; Arte, A.; Sharma, C.; Greenup, R.A.; Reynolds, K.L.; Smith, B.L.; Alexander, B.M.; Moy, B.; Isaikoff, S.J.; et al. Pathological Complete Response after Neoadjuvant Chemotherapy and Impact on Breast Cancer Recurrence and Survival: A Comprehensive Meta-Analysis. Clin. Cancer Res. 2020, 26, 2838–2848. [CrossRef]
10. Kroschinsky, F.; Stólzel, F.; von Bonin, S.; Beutel, G.; Kochanek, M.; Kiehl, M.; Schellongowski, P.; Intensive Care in Hematological and Oncological Patients (iCHOP) Collaborative Group. New Drugs, New Toxicities: Severe Side Effects of Modern Targeted and Immunotherapy of Cancer and Their Management. Crit. Care 2017, 21, 89. [CrossRef]
11. García-Figueiras, R.; Baleato-González, S.; Luna, A.; Muñoz-Iglesias, J.; Oleaga, L.; Vallejo Casas, J.A.; Martín-Noguerol, T.; Broncano, J.; Arrese, M.C.; Vilanova, J.C. Assessing Immunotherapy with Functional and Molecular Imaging and Radiomics. Radiographics 2020, 40, 1987–2010. [CrossRef]
12. Schwartz, L.H.; Litière, S.; de Vries, E.; Ford, R.; Gwyther, S.; Mandrekar, S.; Shankar, L.; Bogaerts, J.; Chen, A.; Dancey, J.; et al. RECIST 1.1-Update and Clarification: From the RECIST Committee. Eur. J. Cancer 2016, 62, 132–137. [CrossRef] [PubMed]
13. Chiou, V.L.; Buratto, M. Pseudoprogression and Immune-Related Response in Solid Tumors. *J. Clin. Oncol.* 2015, 33, 3541–3543. [CrossRef] [PubMed]

14. Ma, Y.; Wang, Q.; Dong, Q.; Zhan, L.; Zhang, J. How to Differentiate Pseudoprogression from True Progression in Cancer Patients Treated with Immunotherapy. *Am. J. Cancer Res.* 2019, 9, 1546–1553. [PubMed]

15. Jia, W.; Gao, Q.; Han, A.; Zhu, H.; Yu, J. The Potential Mechanism, Recognition and Clinical Significance of Tumor Pseudoprogression after Immunotherapy. *Cancer Biol. Med.* 2019, 16, 655–670. [CrossRef]

16. Partridge, S.C.; Zhang, Z.; Newitt, D.C.; Gibbs, J.E.; Chenevert, T.L.; Rosen, M.A.; Bolan, P.J.; Marques, H.S.; Romanoff, J.; Cinimo, L.; et al. Diffusion-Weighted MRI Findings Predict Pathologic Response in Neoadjuvant Treatment of Breast Cancer: The ACRIN 6698 Multicenter Trial. *Radiology* 2018, 289, 618–627. [CrossRef]

17. Li, W.; Newitt, D.C.; Wilmes, L.J.; Jones, E.F.; Arasu, V.; Gibbs, J.; La Yun, B.; Li, E.; Partridge, S.C.; Kornak, J.; et al. Additive Value of Diffusion-Weighted MRI in the I-SPY 2 TRIAL. *J. Magn. Reson. Imaging* 2019, 50, 1742–1753. [CrossRef]

18. Hylton, N.M.; Blume, J.D.; Berreuter, W.K.; Pisano, E.D.; Rosen, M.A.; Morris, E.A.; Weatherall, P.T.; Lehman, C.D.; Newstead, G.M.; Polin, S.; et al. Locally Advanced Breast Cancer: MR Imaging for Prediction of Response to Neoadjuvant Chemotherapy—Results from ACRIN 6657/I-SPY TRIAL. *Radiology* 2012, 263, 663–672. [CrossRef]

19. Hylton, N.M.; Gatsonis, C.A.; Rosen, M.A.; Lehman, C.D.; Newitt, D.C.; Partridge, S.C.; Berreuter, W.K.; Pisano, E.D.; Morris, E.A.; Weatherall, P.T.; et al. Neoadjuvant Chemotherapy for Breast Cancer: Functional Tumor Volume by MR Imaging Predicts Recurrence-Free Survival—Results from the ACRIN 6657/CALGB 150007 I-SPY 1 TRIAL. *Radiology* 2016, 279, 44–55. [CrossRef]

20. Barker, A.D.; Sigman, C.C.; Kelloff, G.J.; Hylton, N.M.; Berry, D.A.; Esserman, L.J. I-SPY 2: An Adaptive Breast Cancer Trial Design in the Setting of Neoadjuvant Chemotherapy. *Clin. Pharmacol. Ther.* 2009, 86, 97–100. [CrossRef]

21. Nanda, R.; Liu, M.C.; Yau, C.; Shatsky, R.; Pusztai, L.; Wallace, A.; Chien, A.J.; Forero-Torres, A.; Ellis, E.; Han, H.; et al. Effect of Pembrolizumab Plus Neoadjuvant Chemotherapy on Pathologic Complete Response in Women with Early-Stage Breast Cancer. *JAMA Oncol.* 2020, 6, 676–684. [CrossRef]

22. Le, N.N.; Li, W.; Onishi, N.; Newitt, D.C.; Gibbs, J.E.; Wilmes, L.J.; Partridge, S.C.; LeStage, B.; Price, E.R.; et al. Effect of Inter-Reader Variability on Diffusion-Weighted MRI Apparent Diffusion Coefficient Measurements and Prediction of Pathologic Complete Response for Breast Cancer. *Tomography* 2022, 8, 1208–1220. [CrossRef] [PubMed]

23. Partridge, S.C.; Nissan, N.; Rahbar, H.; Kitsch, A.E.; Sigmund, E.E. Diffusion-Weighted Breast MRI: Clinical Applications and Emerging Techniques. *J. Magn. Reson. Imaging* 2017, 45, 337–355. [CrossRef]

24. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* 1988, 44, 837–845. [CrossRef] [PubMed]

25. Wilmes, L.J.; McLaughlin, R.L.; Newitt, D.C.; Singer, L.; Sinha, S.P.; Proctor, E.; Wisner, D.J.; Saritas, E.U.; Kornak, J.; Shankaranarayanan, A.; et al. High-Resolution Diffusion-Weighted Imaging for Monitoring Breast Cancer Treatment Response. *Acad. Radiol.* 2013, 20, 581–589. [CrossRef] [PubMed]

26. Minarikova, L.; Bogner, W.; Pinker, K.; Valković, L.; Zaric, O.; Bago-Horvath, Z.; Bartsch, R.; Helbich, T.H.; Trattning, S.; Gruber, S. Investigating the Prediction Value of Multiparametric Magnetic Resonance Imaging at 3 T in Response to Neoadjuvant Chemotherapy in Breast Cancer. *Eur. Radiol.* 2017, 27, 1901–1911. [CrossRef]

27. Li, X.-R.; Cheng, L.-Q.; Liu, M.; Zhang, Y.-J.; Wang, J.-D.; Zhang, A.-L.; Song, X.; Li, J.; Zheng, Y.-Q.; Liu, L. DW-MRI ADC Values Can Predict Treatment Response in Patients with Locally Advanced Breast Cancer Undergoing Neoadjuvant Chemotherapy. *Med. Oncol.* 2012, 29, 425–431. [CrossRef]

28. Park, S.H.; Moon, W.K.; Cho, N.; Song, I.C.; Chang, J.M.; Park, I.-A.; Han, W.; Noh, D.-Y. Diffusion-Weighted MR Imaging: Pretreatment Prediction of Response to Neoadjuvant Chemotherapy in Patients with Breast Cancer. *Radiology* 2010, 257, 56–63. [CrossRef]

29. Song, Y.S.; Choi, S.H.; Park, C.-K.; Yi, K.S.; Lee, W.J.; Yun, T.J.; Kim, T.M.; Lee, S.-H.; Kim, J.-H.; Sohn, C.-H.; et al. True Progression versus Pseudoprogression in the Treatment of Glioblastomas: A Comparison Study of Normalized Cerebral Blood Volume and Apparent Diffusion Coefficient by Histogram Analysis. *Korean J. Radiol.* 2013, 14, 662–672. [CrossRef]

30. Kyriazi, S.; Collins, D.J.; Messiou, C.; Pennert, K.; Davidson, R.L.; Giles, S.L.; Kaye, S.B.; DeSouza, N.M. Metastatic Ovarian and Primary Peritoneal Cancer: Assessing Chemotherapy Response with Diffusion-Weighted MR Imaging—Value of Histogram Analysis of Apparent Diffusion Coefficients. *Radiology* 2011, 261, 182–192. [CrossRef]

31. Chai, L.F.; Prince, E.; Pillarisetty, V.G.; Katz, S.C. Challenges in Assessing Solid Tumor Responses to Immunotherapy. *Cancer Gene Ther.* 2020, 27, 528–538. [CrossRef] [PubMed]

32. Kwak, J.J.; Tirumani, S.H.; Van den Ababee, A.D.; Koo, P.J.; Jacene, H.A. Cancer Immunotherapy: Imaging Assessment of Novel Treatment Response Patterns and Immune-Related Adverse Events. *RadioGraphics* 2015, 35, 424–437. [CrossRef] [PubMed]