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[1]

Nola M. Hylton, PhD
Constantine A. Gatsonis, PhD
Mark A. Rosen, MD, PhD
Constance D. Lehman, MD, PhD
David C. Newitt, PhD
Savannah C. Partridge, PhD
Wanda K. Bernreuter, MD
Etta D. Pisano, MD
Elizabeth A. Morris, MD
Paul T. Weatherall, PhD
Sandra M. Polin, MD
Gillian M. Newstead, MD
Helga S. Marques, MS
Laura J. Esserman, MD, MBA
Mitchell D. Schnall, MD, PhD
For the ACRIN 6657 Trial Team and I-SPY 1 TRIAL Investigators

Purpose: To evaluate volumetric magnetic resonance (MR) imaging for predicting recurrence-free survival (RFS) after neoadjuvant chemotherapy (NACT) of breast cancer and to consider its predictive performance relative to pathologic complete response (PCR).

Materials and Methods: This HIPAA-compliant prospective multicenter study was approved by institutional review boards with written informed consent. Women with breast tumors 3 cm or larger scheduled for NACT underwent dynamic contrast-enhanced MR imaging before treatment (examination 1), after one cycle (examination 2), midtherapy (examination 3), and before surgery (examination 4). Functional tumor volume (FTV), computed from MR images by using enhancement thresholds, and change from baseline (ΔFTV) were measured after one cycle and before surgery. Association of RFS with FTV was assessed by Cox regression and compared with association of RFS with PCR and residual cancer burden (RCB), while controlling for age, race, and hormone receptor (HR)/ human epidermal growth factor receptor type 2 (HER2) status. Predictive performance of models was evaluated by *C* statistics.

Results: Female patients (n = 162) with FTV and RFS were included. At univariate analysis, FTV_2, FTV_4, and ΔFTV_4 had significant association with RFS, as did HR/HER2 status and RCB class. PCR approached significance at univariate analysis and was not significant at multivariate analysis. At univariate analysis, FTV_2 and RCB class had the strongest predictive performance (*C* statistic = 0.67; 95% confidence interval [CI]: 0.58, 0.76), greater than for FTV_4 (0.64; 95% CI: 0.53, 0.74) and PCR (0.57; 95% CI: 0.39, 0.74). At multivariate analysis, a model with FTV_2, ΔFTV_2, RCB class, HR/HER2 status, age, and race had the highest *C* statistic (0.72; 95% CI: 0.60, 0.84).

Conclusion: Breast tumor FTV measured by MR imaging is a strong predictor of RFS, even in the presence of PCR and RCB class. Models combining MR imaging, histopathology, and breast cancer subtype demonstrated the strongest predictive performance in this study.

Online supplemental material is available for this article.

[1]From the Departments of Radiology (N.M.H., D.C.N.) and Surgery (L.J.E.), University of California, San Francisco, 1600 Divisadero St, Room C250, Box 1667, San Francisco, CA 94115; Department of Biostatistics (C.A.G.) and Center for Statistical Sciences (C.A.G., H.S.M.), Brown University, Providence, RI; American College of Radiology Imaging Network (ACRIN), Philadelphia, Pa (C.A.G., H.S.M., M.D.S.); Department of Radiology, Hospital of the University of Pennsylvania, Philadelphia, Pa (M.A.R., M.D.S.); Department of Radiology, University of Washington, Seattle, Wash (C.D.L., S.C.P.); Department of Radiology, University of Alabama, Birmingham, Ala (W.K.B.); Department of Radiology, Medical College of South Carolina, Charleston, SC (E.D.P.); Department of Radiology, Memorial Sloan-Kettering Cancer Center, New York, NY (E.A.M.); Department of Radiology, University of Texas Southwestern Medical Center, Dallas, Tex (P.T.W.); Department of Radiology, Georgetown University, Washington, DC (S.M.P.); and Department of Radiology, University of Chicago, Chicago, Ill (G.M.N.). Received January 20, 2015; revision requested March 10; revision received July 26; accepted August 19; final version accepted September 8. Supported by the American College of Radiology Imaging Network (grants CA079778 and CA080098) and Cancer and Leukemia Group B (grants CA31964 and CA33601). **Address correspondence to** N.M.H. (e-mail: *Nola.Hylton@ucsf.edu*).

Breast cancer is increasingly understood to be a heterogeneous disease that requires tailored treatment strategies. The development of targeted therapies relies on the availability of biomarkers that can be used to assess the effectiveness of promising agents as early and accurately as possible. Imaging has the potential to provide in vivo biomarkers of response that can facilitate the development of new therapeutics. In the area of breast cancer, new treatments are being evaluated in the neoadjuvant setting more frequently when the primary tumor is intact and its response to treatment can be monitored. Magnetic resonance (MR) imaging is particularly effective for demonstrating extent of disease in the breast and is used in the neoadjuvant setting to measure the primary tumor response to treatment. MR imaging showed greater accuracy for prediction of residual disease after neoadjuvant chemotherapy than clinical examination or conventional imaging (1–13).

American College of Radiology Imaging Network (ACRIN) trial 6657 was a multicenter study of contrast material–enhanced MR imaging to assess breast tumor response to neoadjuvant chemotherapy, and the imaging component of the I-SPY TRIAL (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis) (14). ACRIN 6657 was conducted to evaluate MR imaging for measuring response of primary breast cancer to chemotherapy and to assess the benefit of MR imaging for predicting recurrence-free survival (RFS). A previous analysis (15) showed that MR imaging was more strongly associated with pathologic response after neoadjuvant chemotherapy than clinical examination, with the greatest advantage measured early in treatment by using a volumetric measurement of tumor response. The purpose of this study was to evaluate volumetric MR imaging for predicting RFS after neoadjuvant chemotherapy of breast cancer and to consider its predictive performance relative to pathologic complete response (PCR).

Advances in Knowledge

- Functional tumor volume (FTV) measured by MR imaging predicts recurrence-free survival (RFS) for patients who receive neoadjuvant chemotherapy for breast cancer.
- FTV is a stronger predictor of RFS than pathologic complete response (PCR); in this study, univariate *C* statistics for FTV measured before surgery (FTV_4) and PCR were 0.64 (95% confidence interval [CI]: 0.53, 0.74) and 0.57 (95% CI: 0.39, 0.74), respectively.
- FTV predicts RFS as early as after one cycle of standard anthracycline-based chemotherapy; in this study, FTV measured after one cycle (FTV_2) and FTV_4 had comparable univariate *C* statistic (0.67 [95% CI: 0.58, 0.76] and 0.64 [95% CI: 0.53, 0.74], respectively) and comparable multivariate *C* statistic in models that combined FTV and PCR (0.69 [95% CI: 0.57, 0.81] and 0.69 [95% CI: 0.56, 0.81], respectively).
- Predictive performance is highest when MR imaging is combined with histopathologic analysis, with both variables contributing independently; the highest *C* statistic in this study resulted from the multivariate model that included FTV_2, residual cancer burden (RCB) class, and tumor subtype defined by hormone receptor (HR) and human epidermal growth factor receptor type 2 (HER2) status (0.72 [95% CI: 0.60, 0.84]).
- FTV predictive performance differs among breast cancer subtypes defined by HR and HER2 status: exploratory Kaplan-Meier analyses found significant survival differences in the HR-positive/HER2-negative and HER2-positive subtypes by using groups dichotomized by FTV_4 ($P < .001$ and $P = .01$, respectively) but not in the HR-negative/HER2-negative (triple-negative) subtype ($P = .22$); however, significant survival differences were found in the triple-negative subtype by using groups dichotomized by FTV_2 ($P = .006$), but not in the HR-positive/HER2-negative and HER2-positive subtypes ($P = .64$ and $P = .25$, respectively).

Implications for Patient Care

- Breast MR imaging provides prognostic information about tumor response as early as after one cycle of chemotherapy that can potentially be used to guide treatment.
- Breast MR imaging used in combination with the histopathologic variables PCR or RCB may improve the effectiveness of intermediate end points for assessment of drug efficacy in neoadjuvant clinical trials.
- MR imaging helps to show differences in treatment response among breast cancer subtypes that may lead to tailored imaging approaches with improved predictive performance.

Published online before print
10.1148/radiol.2015150013 **Content codes:** BR MR

Radiology 2016; 279:44–55

Abbreviations:
ACRIN = American College of Radiology Imaging Network
FTV = functional tumor volume
FTV_1 = FTV assessed at baseline
FTV_2 = FTV assessed at early treatment time point
FTV_4 = FTV assessed before surgery
HER2 = human epidermal growth factor receptor type 2
HR = hormone receptor
I-SPY TRIAL = Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis
PCR = pathologic complete response
RCB = residual cancer burden
RFS = recurrence-free survival
ΔFTV_2 = change in FTV_2
ΔFTV_4 = change in FTV_4

Author contributions:
Guarantors of integrity of entire study, N.M.H., L.J.E.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, N.M.H., C.A.G., C.D.L., S.C.P., E.A.M., L.J.E.; clinical studies, N.M.H., C.A.G., C.D.L., D.C.N., S.C.P., W.K.B., E.D.P., E.A.M., P.T.W., S.M.P., G.M.N., L.J.E., M.D.S.; experimental studies, P.T.W., M.D.S.; statistical analysis, N.M.H., C.A.G., H.S.M.; and manuscript editing, N.M.H., C.A.G., M.A.R., C.D.L., D.C.N., S.C.P., W.K.B., E.D.P., E.A.M., P.T.W., G.M.N., H.S.M., L.J.E., M.D.S.

Funding:
This research was supported by the National Institutes of Health (grants R01 CA069587 and CA116182).

Conflicts of interest are listed at the end of this article.

Materials and Methods

Study Design and Patient Enrollment

Women with breast tumors that measured 3 cm or greater by clinical examination or imaging and who were scheduled to receive anthracycline-based neoadjuvant chemotherapy were eligible to enroll in Cancer and Leukemia Group B study 150007 and ACRIN 6657. Both protocols were approved by institutional review boards, and patients signed a single consent form. MR examinations were performed within 4 weeks before administration of anthracycline-cyclophosphamide chemotherapy (examination 1), at least 2 weeks after the first cycle of anthracycline-cyclophosphamide and before the second cycle of anthracycline-cyclophosphamide (examination 2), after all anthracycline-cyclophosphamide and before taxane if taxane was administered (examination 3), and after completion of chemotherapy and before surgery (examination 4). Mammography was performed at baseline and presurgical time points only, but it was not considered in this analysis. Core biopsies were performed at baseline and between 24 and 96 hours after starting treatment as part of the Cancer and Leukemia Group B 150007 protocol.

MR Imaging Protocol

MR imaging was performed by using a 1.5-T field-strength MR imaging system and a dedicated four- or eight-channel breast radiofrequency coil. Patients were placed on the MR imaging table in the prone position with an intravenous catheter inserted in the antecubital vein or hand. The image acquisition protocol included a localization sequence and T2-weighted sequence followed by a contrast-enhanced T1-weighted series. All imaging was performed unilaterally over the symptomatic breast in the sagittal orientation. The contrast-enhanced series consisted of a high-resolution (≤1 mm in-plane spatial resolution), three-dimensional, fat-suppressed, T1-weighted gradient-echo sequence (repetition time msec/echo time msec, ≤20/4.5; flip angle, ≤45°; field of view, 16–18 cm; minimum matrix, 256 × 192; 64 sections; section thickness, ≤2.5 mm). Imaging time length for the T1-weighted sequence was between 4.5 and 5 minutes. The sequence was performed once before injection of contrast agent and repeated two to four times after injection of the contrast agent. Interim-aging delays were added as needed to result in postcontrast administration temporal sampling between 2 minutes 15 seconds and 2 minutes 30 minutes for early phase images and between 7 minutes 15 seconds and 7 minutes 45 seconds for delayed phase images.

Image Analysis

After each MR examination, image data were transferred to the ACRIN Core Laboratory for central archival and subsequently to the University of California at San Francisco for image analysis. The

Figure 1

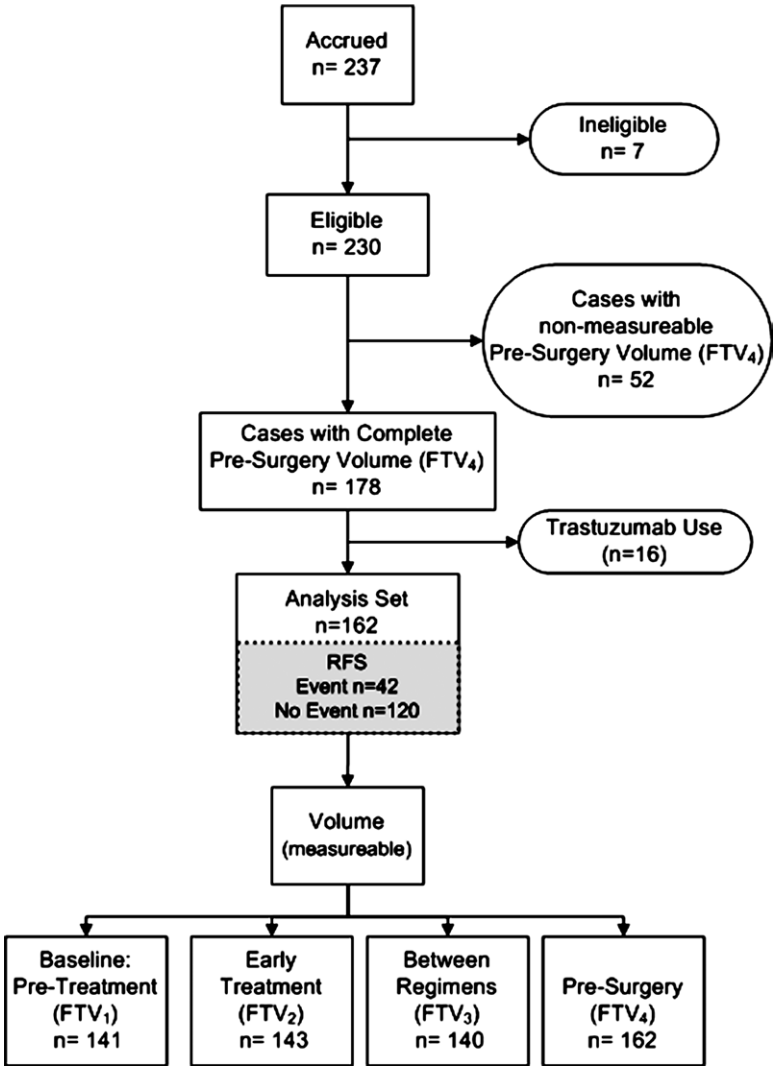

Figure 1: Study flowchart shows MR imaging data used in the analysis set. Of 230 eligible patients, 162 patients with measurable FTV_4 data who did not receive any trastuzumab were included in the analysis. Reduced subsets at FTV_1, FTV_2, and between-regimen *(FTV_3)* time points were because of missing MR examinations or nonmeasurable FTV.

Table 1

Characteristics of Patients in ACRIN 6657

Parameter	Eligible Set ($n = 230$)	Included in Analysis ($n = 162$)	Excluded from Analysis ($n = 68$)	*P* Value
Age				.67*
Median†	49 (26–68)	48 (26–68)	49 (29–68)	…
Mean‡	47.7 ± 8.9	47.6 ± 9.1	48.0 ± 8.6	…
Race				.27
Asian	9 (3.9)	7 (4.3)	2 (2.9)	…
Black or African American	46 (20.0)	29 (17.9)	17 (25.0)	…
White	170 (73.9)	124 (76.5)	46 (67.6)	…
More than one race	1 (0.4)	0 (0.0)	1 (1.5)	…
Unknown	4 (1.7)	2 (1.2)	2 (2.9)	…
Ethnicity				.55
Hispanic or Latino	10 (4.3)	7 (4.3)	3 (4.4)	…
Not Hispanic or Latino	203 (88.3)	145 (89.5)	58 (85.3)	…
Unknown	17 (7.4)	10 (6.2)	7 (10.3)	…
Menopausal status				.55
Before menopause	106 (46.1)	77 (47.5)	29 (42.6)	…
After menopause	76 (33.0)	59 (36.4)	17 (25.0)	…
Indeterminate	38 (16.5)	26 (16.0)	12 (17.6)	…
Missing	10 (4.3)	0 (0.0)	10 (14.7)	…
Single or multiple lesions				.10
Single	130 (56.5)	99 (61.1)	31 (45.6)	…
Multiple	95 (41.3)	63 (38.9)	32 (47.1)	…
Missing	5 (2.2)	0 (0.0)	5 (7.4)	…
Axillary lymph node status at initial staging				.12
No	86 (37.4)	59 (36.4)	27 (39.7)	
Yes	132 (57.4)	103 (63.6)	29 (42.6)	
Missing	12 (5.2)	0 (0.0)	12 (17.6)	
HR status				.48
Negative	87 (37.8)	62 (38.3)	25 (36.8)	…
Positive	131 (57.0)	99 (61.1)	32 (47.1)	…
Missing	12 (5.2)	1 (0.6)	11 (16.2)	…
HER2 status				.01
Negative	149 (64.8)	118 (72.8)	31 (45.6)	…
Positive	66 (28.7)	41 (25.3)	25 (36.8)	…
Missing	15 (6.5)	3 (1.9)	12 (17.6)	…
HR/HER2				.02
HR-negative/ HER2-negative (triple negative)	53 (23.0)	40 (24.7)	13 (19.1)	…
HR-positive/ HER2-negative	96 (41.7)	78 (48.1)	18 (26.5)	…
HER2-positive	66 (28.7)	41 (25.3)	25 (36.8)	…
Missing	15 (6.5)	3 (1.9)	12 (17.6)	…
Pretreatment invasive histology analysis				.87
Ductal carcinoma	175 (76.1)	130 (80.2)	45 (66.2)	…
Lobular carcinoma	18 (7.8)	14 (8.6)	4 (5.9)	…
Mixed ductal/lobular carcinoma	8 (3.5)	5 (3.1)	3 (4.4)	…
Other	16 (7.0)	12 (7.4)	4 (5.9)	…
Missing	13 (5.7)	1 (0.6)	12 (17.6)	…

Table 1 (continues)

primary predictor variable (functional tumor volume [FTV]) was measured by semiautomated computer analysis of contrast-enhanced images by using the signal enhancement ratio method (16). FTV was then automatically computed by summing all image voxels within the volume of interest by meeting thresholds for percent enhancement and signal enhancement ratio and a minimum number of connected pixels above threshold. Percent enhancement threshold was nominally set at 70% and adjusted as needed for each site in order to account for variability in MR imaging systems and imaging parameters. Adjusted percent enhancement threshold levels were determined empirically for each site by the research associate on the basis of visual agreement of computer-identified and qualitatively assessed extent of tumor. All MR imaging studies from a given site were processed by using the same site-specific percent enhancement threshold. Signal enhancement ratio threshold was set to 0 for total FTV measurement. FTV was assessed at baseline (FTV_1), the early treatment time point (FTV_2), and before surgery (FTV_4). Change in FTV_2 or FTV_4 (ΔFTV_2 and ΔFTV_4, respectively) was defined relative to baseline, (ie, $\Delta FTV_2 = FTV_2 - FTV_1$; and $\Delta FTV_4 = FTV_4 - FTV_1$).

Assessment of Response and RFS End Points

Tumor diameter in centimeters was estimated by physical examination at all time points for palpable tumors. Location in the breast was also recorded. Clinical response was assessed at time points after baseline according to Response Evaluation Criteria in Solid Tumors version 1.1 (known as RECIST) as complete response, partial response, stable disease, or progressive disease (17).

Standard histopathologic analysis was performed at study sites by the institutional pathologist according to the I-SPY TRIAL protocol (14,18). PCR was defined as no residual invasive disease in either breast or axillary lymph nodes after neoadjuvant therapy. A subsequent central pathologic review was conducted by one pathologist to

Table 1 (continued)

Characteristics of Patients in ACRIN 6657

Parameter	Eligible Set (n = 230)	Included in Analysis (n = 162)	Excluded from Analysis (n = 68)	P Value
Pretreatment histologic grade				.46
Grade I	18 (7.8)	12 (7.4)	6 (8.8)	…
Grade II	95 (41.3)	75 (46.3)	20 (29.4)	…
Grade III	103 (44.8)	72 (44.4)	31 (45.6)	…
Indeterminate	4 (1.7)	3 (1.9)	1 (1.5)	…
Missing	10 (4.3)	0 (0.0)	10 (14.7)	…
Pretreatment clinical size (cm)				.21
Median†	6 (2.0–99.0)	6 (2.0–99.0)	6 (2.0–99.0)	…
25th percentile	4.0	4.0	4.7	…
75th percentile	8.0	8.0	8.0	…
95th percentile	13.0	11.0	25.0	…
Mean‡	7.4 ± 9.5	6.8 ± 7.8	8.9 ± 13.1	…
Early treatment volume (cm^3)				.77
Median†	8.9 (0–177.6)	8.9 (0–177.6)	8.7 (0.2–71.9)	…
25th percentile	2.9	3.0	2.0	…
75th percentile	22.8	24.2	18.4	…
95th percentile	63.0	63.0	53.1	…
Mean‡	16.9 ± 22.9	17.5 ± 24.17	14.5 ± 16.9	…
Presurgery volume (cm^3)				.30
Median†	0.11 (0–73.1)	0.12 (0–73.1)	0.04 (0–66.24)	…
25th percentile	0	0	0	…
75th percentile	0.97	1.10	0.36	…
95th percentile	8.52	8.28	66.24	…
Mean‡	2.4 ± 8.9	2.2 ± 7.8	4.7 ± 16.5	…
MR imaging time points performed§				
Baseline	225 (97.8)	162 (100.0)	63 (92.6)	
Early treatment	213 (92.6)	156 (96.3)	57 (83.8)	
Between regimens	202 (87.8)	153 (94.4)	49 (72.1)	
Before surgery	211 (91.7)	162 (100.0)	49 (72.1)	
Presurgery clinical response				.06
Complete response	90 (39.1)	67 (41.4)	23 (33.8)	…
Partial response	92 (40.0)	72 (44.4)	20 (29.4)	…
Stable disease	18 (7.8)	10 (6.2)	8 (11.8)	…
Progressive disease	7 (3.0)	3 (1.9)	4 (5.9)	…
Missing	23 (10.0)	10 (6.2)	13 (19.1)	…
In situ present				.25
No	102 (44.3)	79 (48.8)	23 (33.8)	…
Yes	112 (48.7)	79 (48.8)	33 (48.5)	…
No surgery	6 (2.6)	4 (2.5)	2 (2.9)	…
Missing	10 (4.3)	0 (0.0)	10 (14.7)	…
RCB class‖				.92
0	56 (24.3)	41 (25.3)	15 (22.1)	…
I	17 (7.4)	13 (8.0)	4 (5.9)	…
II	86 (37.4)	67 (41.4)	19 (27.9)	…
III	41 (17.8)	32 (19.8)	9 (13.2)	…
Missing	30 (13.0)	9 (5.6)	21 (30.9)	…

Table 1 (continues)

estimate residual cancer burden (RCB), a histopathologic measurement that reflects the size and cellularity of residual disease in the breast and axillary lymph nodes (19). RCB was measured as a continuous variable (RCB index) and further categorized into ranges reflecting no remaining invasive disease (RCB class 0) and increasing amounts of residual disease (RCB class I, II, or III) according to the methods described by Symmans et al (19).

Hormone receptor (HR) positivity (estrogen receptor–positive or partial response–positive) and human epidermal growth factor receptor type 2 (HER2) expression were determined from pretreatment core biopsy by immunohistochemistry and Allred score. For subtype analysis, patient groups were defined as those with tumors that were HR-positive/HER2-negative, HER2-positive, and HR-negative/HER2-negative (triple negative).

RFS was defined as the time between first chemotherapy treatment and disease recurrence according to the STEEP criteria (20).

Analysis Set

We enrolled 237 patients between May 2002 and March 2006 at nine institutions. Seven patients were subsequently found ineligible because of medical contraindications, which resulted in 230 eligible patients. For analysis of the primary study variable, FTV_4, 52 additional patients were excluded because examination 4 was not performed (n = 14), examination 4 data were not received from the site (n = 8), or FTV_4 was not measurable because of protocol deviations, poor image quality, or artifacts (n = 30). This yielded a set of 178 patients with FTV_4 measurement. Because neoadjuvant trastuzumab was not used as standard therapy until 2005, only a small number (n = 16) of enrolled patients underwent trastuzumab therapy and they were excluded from the analysis. The resulting analysis set included 162 patients with FTV_4 (Fig 1). Reduced subsets at FTV_1 (n = 141), FTV_2 (n = 143), and between-regimen (n = 140) time points were because of

Table 1 (continued)

Characteristics of Patients in ACRIN 6657

Parameter	Eligible Set (n = 230)	Included in Analysis (n = 162)	Excluded from Analysis (n = 68)	P Value
PCR				.77
Complete responder	58 (25.2)	42 (25.9)	16 (23.5)	…
Nonresponder	156 (67.8)	116 (71.6)	40 (58.8)	…
Missing	16 (7.0)	4 (2.5)	12 (17.6)	…
Total pathologic size (mm)				.84
Median†	14 (0–150)	13 (0–150)	14 (0–150)	…
Mean‡	23.1 ± 29.8	23.0 ± 30.2	23.4 ± 28.8	…
Surgery type				.71
Lumpectomy	91 (39.6)	66 (40.7)	25 (36.8)	…
Mastectomy	123 (53.5)	92 (56.8)	31 (45.6)	…
Missing	16 (7.0)	4 (2.5)	12 (17.6)	…
RFS				.14
Event (first occurrence of local or distant progression or death)	63 (27.4)	42 (25.9)	21 (30.9)	…
No event	157 (68.3)	120 (74.1)	37 (54.4)	…
Missing	10 (4.3)	0 (0.0)	10 (14.7)	…
Local and/or distant progression				.26
Not local or distant	5 (2.2)	3 (1.9)	2 (2.9)	…
Local only	6 (2.6)	3 (1.9)	3 (4.4)	…
Local (ie, distant missing)	1 (0.4)	1 (0.6)	0 (0.0)	…
Distant only	38 (16.5)	23 (14.2)	15 (22.1)	…
Both local and distant	12 (5.2)	11 (6.8)	1 (1.5)	…
Both missing	11 (4.8)	1 (0.6)	10 (14.7)	…
Not applicable (ie, no event)	157 (68.3)	120 (74.1)	37 (54.4)	…

Note.—Unless otherwise indicated, data are numerators and data in parentheses are percentages.

* Wilcoxon P value was used for continuous variables and χ^2 P value was used for categorical variables.

† Data in parentheses are range.

‡ Data are ± standard deviation.

§ P value excluded because MR imaging was not available.

‖ RCB class definitions: 0, RCB index 0; I, RCB index ≤1.3; II, RCB index 1.36 to ≤3.28; III, RCB index >3.28.

missing MR examinations or nonmeasurable FTV. RFS was available for all cases included in the analysis set. The median follow-up time was 3.9 years (range, 0.5–6.9 years). PCR and RCB were unavailable for four patients who did not undergo surgery. RCB was missing for an additional five patients because histopathologic slides were unavailable for central re-review.

All patients included in the analysis group for this study were included in the set of 216 patients previously reported (15). The previous study evaluated the relationship between FTV and the intermediate histopathologic end points PCR and RCB, whereas in this study we reported on the relationship with the survival outcome RFS.

Statistical Analysis

Cox regression was used to examine the association of RFS to the continuous FTV variables and to compare with the association to PCR status and RCB class. The following patient characteristics and clinical variables were also included in the analysis: age, ethnicity, clinical response status, HR status, HER2 status, and HR/HER2 status. Univariate and multivariate models were fitted and compared by using likelihood ratio tests. To assess the predictive performance of models, a c-statistic modified to account for censoring was computed, together with a 95% confidence interval (21). Overall model fit (calibration) was assessed by using the Gronnesby-Borgan test (22). In exploratory analysis of the associations of FTV with RFS, Kaplan-Meier curves were computed for RFS by patient groups defined by dichotomizing FTV_2 and FTV_4 at the lowest quartile, median quartile, and highest quartile values. In addition, Kaplan-Meier curves for RFS were computed separately for subgroups defined by HR/HER2 status, with FTV dichotomized at the highest quartile. Survival curves were compared by using a log-rank test. For the comparisons of the subgroups included and excluded from the analysis a Wilcoxon P value was used for continuous variables and χ^2 P value test was used for categorical variables.

Computations were performed by using statistical software (SAS version 9.4, SAS Institute, Cary, NC; Stata version 13, Stata, College Station, Tex; and R version 3.0.2, R Project for Statistical Computing, Vienna, Austria). A P value less than .05 was required to indicate statistical significance to test a single hypothesis.

Results

Patient Characteristics

Patient characteristics are listed in Table 1 and were comparable for the eligible and analysis subsets. Longitudinal MR images and associated FTV maps are shown in Figure 2 for a patient with an excellent treatment response.

Imaging and Biomarker Analysis

In univariate analysis of 13 variables, FTV_2 and FTV_4 had the smallest P values for association with RFS (both $P < .001$; Table 2). ΔFTV_4 ($P = .006$), HER2 status ($P = .02$), HER2-positive versus HR-positive/HER2-negative subtype ($P = .006$), and RCB class III versus 0 ($P = .002$) also showed significant association with RFS. HR, PCR, age, race, and clinical response were not significantly associated with RFS.

Figure 2

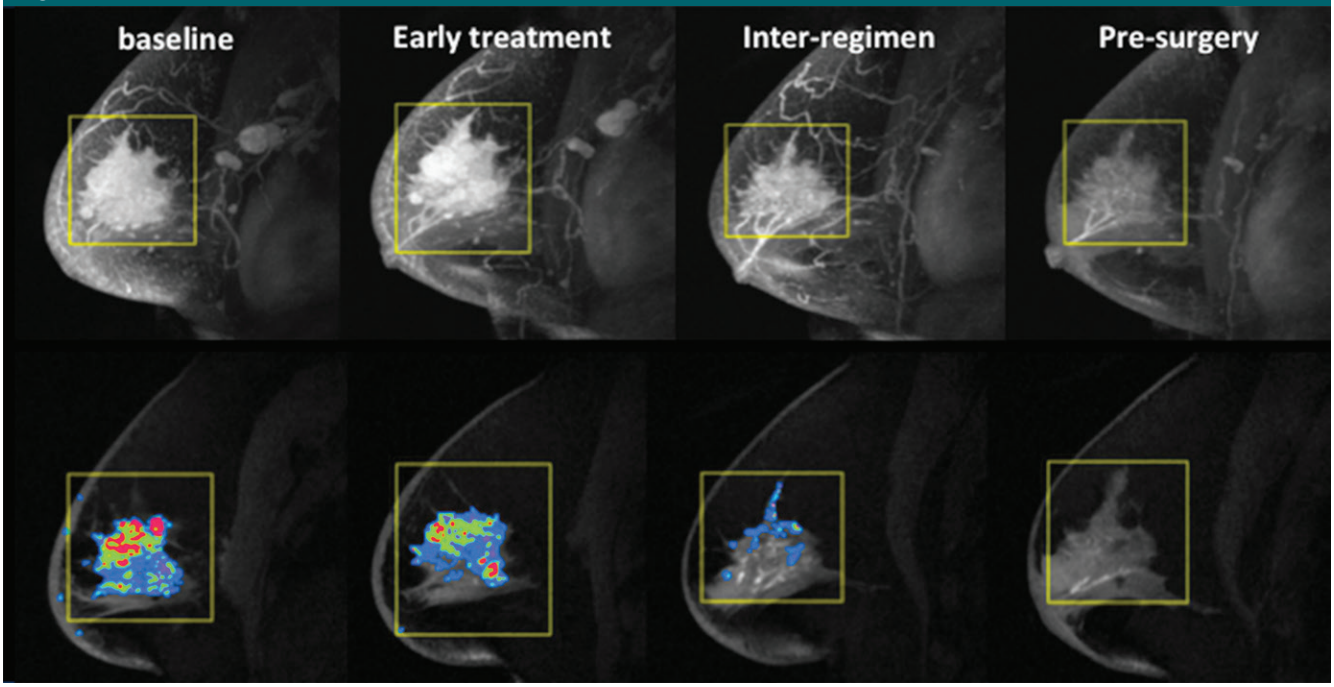

Figure 2: Longitudinal MR images and FTV maps. Maximum intensity projection images (top row) and corresponding FTV maps (bottom row) for a patient with an excellent clinical response and disseminated residual disease. FTV measurements were 48.5 cm^3, 35.4 cm^3, 5.6 cm^3, and 0 cm^3 for baseline, early treatment, inter-regimen, and presurgery time points, respectively (shown from left to right).

Multivariate models, which included volume variables at the presurgical time point, FTV_4, and ΔFTV_4 combined with PCR and HR/HER2 subtype, are shown in Table 3. The volume variables remain significant in all models, whereas PCR was not significant when added to a model with FTV_4 and ΔFTV_4. HR/HER2 status was significant when added to a model with presurgery volume variables and PCR. When FTV_4 and ΔFTV_4 were replaced with FTV_2 and ΔFTV_2 in the three models, FTV_2 was the only statistically significant contributor (Table E1 [online]). In similar analyses that replaced PCR with RCB class, RCB class III versus 0 did add significantly to models 2 and 3 (Tables E2, E3 [online]).

The confidence intervals for univariate and multivariate models overlap extensively, which indicates that there were no statistically significant differences between models (Table 4). In univariate modeling, the highest *C* statistic was found for FTV_2 (0.67 [95% confidence interval: 0.58, 0.76]) and RCB class (0.67 [95% confidence interval: 0.54, 0.81]). The univariate *C* statistic for PCR was 0.57 (95% confidence interval: 0.39, 0.74). Multivariate models resulted in higher *C* statistics in general, and the highest values were found by using the early-treatment FTV_2 measurements combined with RCB class. The model that combined FTV_2, ΔFTV_2, RCB class, and HR/HER2 subtype had the highest *C* statistics (0.72 [95% confidence interval: 0.60, 0.84]).

Kaplan-Meier plots that compared RFS estimates by using FTV cut points at lowest quartile, median quartile, and highest quartile are shown for FTV_2 and FTV_4 in Figure 3. The corresponding plots by HR/HER2 status are shown in Figure 4. For both time points, the greatest separation was found by using highest quartile, separating patients in the quartile with the most residual FTV from all others (Fig 3).

Discussion

In this study, breast tumor volume measured with MR imaging before and during neoadjuvant chemotherapy was found to predict RFS. These findings confirm previous single-institution results (11) that show association of MR imaging–measured breast tumor volume with RFS. At univariate analysis, the early time-point measurement FTV_2 and the presurgical measurements FTV_4 and ΔFTV_4 were strongly associated with RFS, as was RCB class. PCR did not reach statistical significance in this analysis, although in an analysis of 172 patients in the I-SPY TRIAL that encompassed the subset included here, PCR was found to be statistically significant (18). FTV_2, measured after only one cycle of anthracycline chemotherapy, had the highest univariate *C* statistic, comparable to RCB class, which suggested that MR imaging can provide prognostic information early in treatment.

Predictive performance was improved in multivariate models that combined MR imaging and histopathologic variables. Multivariate models that used the early time-point variables

Table 2

Univariate Cox Regression Analysis for Association with RFS

Parameter	Hazard Ratio	*P* Value
Age	0.99 (0.96, 1.02)	.57
Race (African American vs white)	1.03 (0.46, 2.34)	.94
Race (other vs white)	0.88 (0.21, 3.68)	.86
HR (negative vs positive)	1.74 (0.94, 3.21)	.08
HER2 (positive vs negative)	2.22 (1.17, 4.20)	.02
Subtype (HR-negative/HER2-negative [triple negative] vs HR-positive/HER2-negative)	1.85 (0.83, 4.14)	.13
Subtype (HER2-positive vs HR-positive/HER2-negative)	2.82 (1.35, 5.90)	.006
Early treatment		
FTV_2	1.02 (1.01, 1.04)	<.001
ΔFTV_2	0.99 (0.98, 1.00)	.08
Presurgery		
FTV_4	1.07 (1.04, 1.10)	<.001
ΔFTV_4	0.99 (0.98, 1.00)	.006
Clinical response (partial response vs complete response)	0.86 (0.45, 1.66)	.66
Clinical response (stable disease vs complete response)	0.72 (0.20, 2.58)	.62
Clinical response (progressive disease vs complete response)	1.62 (0.22, 12.19)	.64
Postsurgical pathologic analysis		
PCR (nonresponder vs responder)	2.02 (0.84, 4.83)	.12
RCB class I versus 0	0.51 (0.06, 4.26)	.54
RCB class II versus 0	1.36 (0.52, 3.59)	.53
RCB class III versus 0	4.49 (1.76, 11.50)	.002
RCB index	1.72 (1.31, 2.24)	<.001

Note.—Data in parentheses are 95% confidence intervals.

FTV_2 and ΔFTV_2 performed slightly better than similar models that used the presurgical variables FTV_4 and ΔFTV_4, but this was not statistically significant. Improvements in predictive performance were also gained when PCR was replaced in multivariate models by the more refined system of RCB class for assessing residual disease. The highest overall *C* statistic was obtained when early time-point information from FTV_2 was combined with RCB class, and both variables were statistically significant. While multivariate models were adjusted for age and ethnicity, they did not consider other potential clinically relevant predictors, such as comorbidity, past history of screening, or bone density, which might have improved prediction.

The observation that the best prediction was obtained when imaging and histopathologic analysis were combined may reflect the complementary strengths of histopathologic analysis and MR imaging to detect and measure residual disease when little disease remains and when extensive disease remains, respectively. While histopathologic analysis can accurately assess the

Table 3

Multivariate Cox Regression Analysis Including FTV_4, PCR, and HR/HER2 Subtype

	Model 1: FTV_4 and ΔFTV_4		Model 2: FTV_4, ΔFTV_4, and PCR		Model 3: FTV_4, ΔFTV_4, PCR, and HR/HER2	
Parameter	Hazard Ratio	*P* Value	Hazard Ratio	*P* Value	Hazard Ratio	*P* Value
FTV_4	1.08 (1.05, 1.11)	<.001	1.08 (1.05, 1.11)	<.001	1.08 (1.05, 1.12)	<.001
ΔFTV_4	0.99 (0.98, 0.99)	.001	0.99 (0.98, 0.99)	.001	0.99 (0.98, 1.00)	.005
PCR						
Complete response						
Nonresponder			1.46 (0.59, 3.65)	.42	1.78 (0.69, 4.55)	.23
HR/HER2						
HR-positive/ HER2-negative						
HR-negative/ HER2-negative					1.86 (0.74, 4.64)	.18
HER2-positive					2.97 (1.27, 6.96)	.01
Age (y)	1.00 (0.97, 1.04)	.85	0.99 (0.96, 1.03)	.78	1.00 (0.96, 1.05)	.91
Race						
White						
African American	0.84 (0.36, 2.00)	.70	0.86 (0.36, 2.05)	.73	0.80 (0.32, 1.95)	.62
Other	1.10 (0.25, 4.80)	.90	1.15 (0.26, 4.99)	.86	1.15 (0.27, 4.97)	.85
Modified *C* statistic	0.68 (0.57, 0.79)		0.69 (0.56, 0.81)		0.67 (0.54, 0.80)	

Note.—Data in parentheses are 95% confidence interval range.

Table 4

Table of *C* Statistics and 95% Confidence Intervals for Association with RFS

Model	*C*-Statistic	Standard Error
Univariate		
PCR	0.57 (0.39, 0.74)	0.09
RCB class	0.67 (0.54, 0.81)	0.07
FTV_2	0.67 (0.58, 0.76)	0.05
ΔFTV_2	0.54 (0.42, 0.66)	0.06
FTV_4	0.64 (0.53, 0.74)	0.06
ΔFTV_4	0.62 (0.50, 0.73)	0.06
Age	0.53 (0.40, 0.65)	0.07
Race	0.51 (0.37, 0.65)	0.07
HR	0.57 (0.40, 0.74)	0.09
HER2	0.59 (0.42, 0.75)	0.08
HR/HER2	0.55 (0.39, 0.72)	0.08
Clinical response	0.52 (0.38, 0.65)	0.07
Multivariate		
FTV_2 and ΔFTV_2	0.70 (0.60, 0.80)	0.05
FTV_2, ΔFTV_2, and PCR	0.69 (0.57, 0.81)	0.06
FTV_2, ΔFTV_2, and PCR , and HR/HER2	0.69 (0.58, 0.81)	0.06
FTV_2, ΔFTV_2, and RCB class	0.72 (0.60, 0.84)	0.06
FTV_2, ΔFTV_2, RCB class, and HR/HER2	0.72 (0.60, 0.84)	0.06
FTV_4 and ΔFTV_4	0.68 (0.57, 0.79)	0.06
FTV_4, ΔFTV_4, and PCR	0.69 (0.56, 0.81)	0.06
FTV_4, ΔFTV_4, PCR , and HR/HER2	0.67 (0.54, 0.80)	0.07
FTV_4, ΔFTV_4, and RCB class	0.72 (0.60, 0.84)	0.06
FTV_4, ΔFTV_4, RCB class, and HR/HER2	0.72 (0.59, 0.84)	0.06

Note.—Multivariate models were adjusted for age and ethnicity. Data in parentheses are 95% confidence intervals.

presence and size of minimal disease, it may not provide an accurate estimate of extensive residual disease because of tissue distortion from specimen processing and sampling limitations. Conversely, MR imaging is accurate for estimation of the size and extent of substantial residual disease, particularly multifocal cancers, but is less sensitive when minimal disease is present, such as after neoadjuvant chemotherapy (4,23–25).

Interestingly, in both univariate and multivariate analyses, the absolute value of FTV showed stronger association with RFS than change from baseline. This especially applied to the early treatment time point. FTV_2 resulted in the highest overall univariate *C* statistic of 0.67 (95% confidence interval: 0.58, 0.76); however, the *C* statistic for ΔFTV_2 was only 0.54 (95% confidence interval: 0.42, 0.66), and ΔFTV_2 did not contribute significantly to multivariate models. A possible explanation for this observation is the uncertainty in baseline FTV measurements caused by false-positive findings of background parenchymal enhancement that may have contributed to error in response measurements. Background parenchymal enhancement is a common occurrence in premenopausal women and has been shown to be reduced by hormonal or chemotherapy treatment (26–28). Uncertainty in baseline FTV would be expected to contribute more error to early response measurements than those made later in the course of treatment. Accordingly, at the later presurgical time point, both FTV_4 and ΔFTV_4 were significant in univariate analysis and both contributed significantly in multivariate models.

Exploratory Kaplan-Meier analyses supported the findings from univariate and multivariate Cox analysis that showed comparable or better prediction by using early time point versus presurgical measurement. In Kaplan-Meier analysis, FTV_2 was able to discriminate groups with different survival estimates by using any of the three cut points considered, but FTV_4 was only able to discriminate groups by using the highest quartile. The failure of FTV_4 to identify statistically significant survival differences by using the lowest quartile and median quartile cut points was likely because of the greater magnitude of responses at examination 4 and the large number of complete imaging responses that fell below the lowest quartile and median quartile.

The Kaplan-Meier analyses performed by subtype suggest that both the ability of FTV to discriminate differences in survival outcomes and the optimal timing of MR imaging measurement differ by subtype. At the presurgery time point, greater RFS separation was found in the HR-positive/HER2-negative and HER2-positive subtypes than in the whole cohort, but no significant difference was found in the triple-negative subtype. The inability of MR imaging to detect an RFS difference in the triple-negative subtype at stage T4 may be attributable to several factors, including the small number of patients in the subset (40 of 162), the paradoxical characteristic of triple-negative tumors to exhibit excellent chemotherapeutic response despite poor survival outcomes (29), and the poor ability of MR imaging to detect minimum residual disease. However, a significant and large RFS separation was found in the triple-negative group at stage T2, which suggests that the information provided by MR imaging may be maximized earlier for triple-negative tumors than for other tumor subtypes. It is important to note that because the quartile cut points are data dependent and may not be applicable to other cohorts, these Kaplan-Meier analyses are only illustrative. Additional studies are needed to confirm these findings.

Twenty-three percent of eligible participants were excluded from this analysis because of image data that was insufficient and unable to be analyzed. A majority of these exclusions

Figure 3

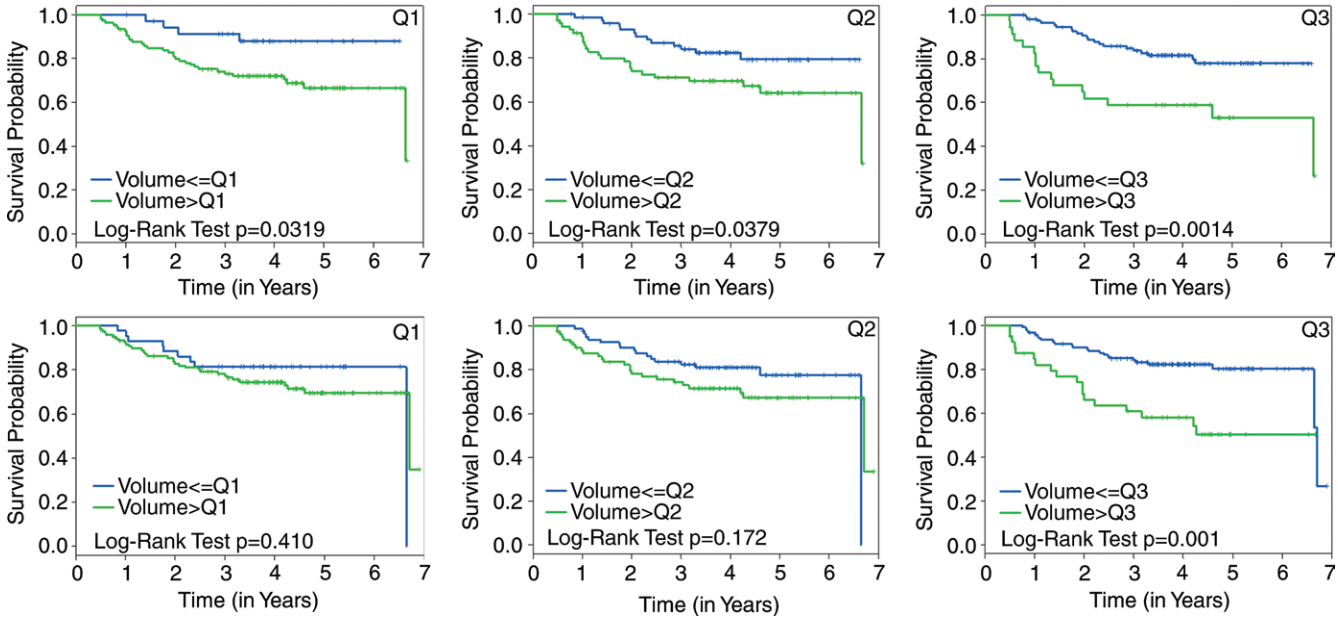

Figure 3: Graphs show Kaplan-Meier RFS estimates according to FTV quartile cut points. RFS stratified by FTV_2 (top) is compared with FTV_4 (bottom) cut points at the lowest quartile *(Q1)*, median quartile *(Q2)*, and highest quartile *(Q3)*, respectively, left to right. The log-rank test *P* value is shown for each plot.

Figure 4

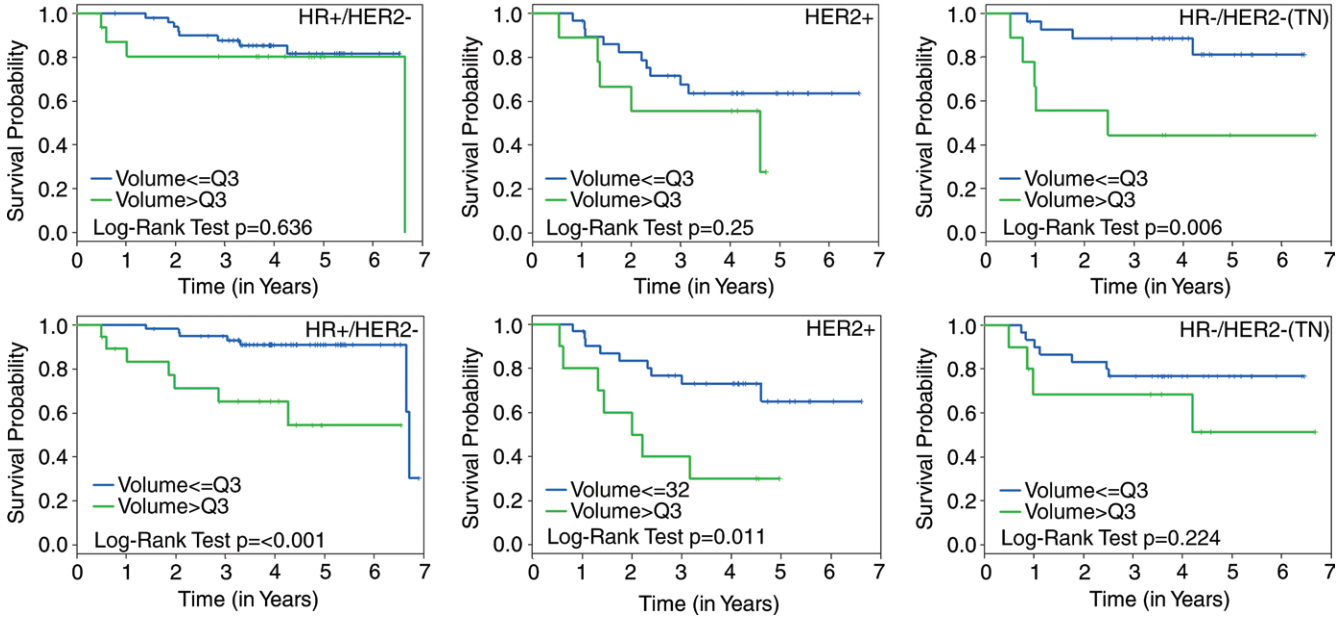

Figure 4: Graphs show Kaplan-Meier plots with RFS estimates by time point and HR and HER2 subtype. RFS stratified by FTV_2 (top row) is compared with FTV_4 (bottom row) by using the highest quartile *(Q3)* cut point for HR-positive *(HR+)* and HER2-negative *(HER2-)*, HER2-positive *(HER2+)*, and HR-negative/HER2-negative (*HR-/HER2-*; triple negative) subtypes, respectively, left to right. The log-rank test *P* value is shown for each plot.

were because of problems that might have been corrected with repeat imaging. The exclusions did not result in substantial differences in characteristics between the eligible and analysis sets, although they may explain the lack of statistical significance of PCR found in this study compared with the parent data set (18). Time delays between examinations, image submission, and data review contributed to the high rate of data loss. Real-time review of protocol compliance and image quality might have allowed more timely correction of protocol deviations and repeat imaging for cases in which inadequate image quality or artifacts prevented FTV measurement. Several quality control procedures have been implemented in the subsequent I-SPY 2 TRIAL to reduce image data loss, including protocol compliance and image data quality review within 72 hours and prior to patient randomization.

ACRIN 6657 was conducted as a prospective multicenter study to assess MR imaging for measuring breast cancer response to neoadjuvant chemotherapy. Results of this study show that MR imaging measurements of breast tumor volume by using a standardized method are effective to predict RFS as early as one cycle of treatment. The imaging studies also suggest that MR imaging captures meaningful differences in response among breast cancer subtypes and reinforces findings from the parent I-SPY TRIAL, which shows that prediction improves in the context of subtype. These studies also suggest that imaging provides information independent from histopathologic analysis. In 2012, the U.S. Food and Drug Administration released a draft guidance regarding the use of PCR as an end point to support accelerated approval of promising new drugs for the neoadjuvant treatment of high-risk early-stage breast cancer (30). Information provided by MR imaging about tumor response may improve the value of PCR as an intermediate end point for RFS. MR imaging has the additional benefit of allowing tumor measurement before surgery and may enable changes to therapy before excision. Experience in ACRIN 6657 led to improvements in the quality and reliability of breast MR imaging performed in the multicenter clinical trials setting. Additional benefits may be realized through more advanced quantitative methods, including diffusion-weighted MR imaging, and these methods are being investigated in continuing trials.

Acknowledgments: The authors sincerely thank the patients who participated in this trial, and the patient advocates for their dedicated support. We also gratefully acknowledge and thank those individuals who contributed substantially to the work reported in this manuscript, including the ACRIN 6657 trial team, the I-SPY TRIAL investigators network, and the staff members who contributed to the conduct of the study at University of California at San Francisco, University of Pennsylvania, University of North Carolina at Chapel Hill, Georgetown University, University of Alabama, Memorial Sloan-Kettering Cancer Center, University of Texas, Southwestern, University of Washington and University of Chicago. The authors also thank Jessica Gibbs, BS, at the University of California San Francisco breast MR imaging laboratory for performing the image analysis.

Disclosures of Conflicts of Interest: N.M.H. disclosed no relevant relationships. **C.A.G.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: author disclosed paid consultancies for Wilex AG, Endocyte, Genentech, Philips Healthcare, and EBG Advisors; author disclosed payment for lectures from Medical Imaging and Technology Alliance. Other relationships: disclosed no relevant relationships. **M.A.R.** disclosed no relevant relationships. **C.D.L.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: author disclosed personal fees from GE and Bayer and a grant from GE. Other relationships: disclosed no relevant relationships. **D.C.N.** disclosed no relevant relationships. **S.C.P.** disclosed no relevant relationships. **W.K.B.** disclosed no relevant relationships. **E.D.P.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: author disclosed Founder and Board Chair of NextRay; and contracts for imaging research from Fuji and Koning Corporation. Other relationships: disclosed no relevant relationships. **E.A.M.** disclosed no relevant relationships. **P.T.W.** disclosed no relevant relationships. **S.M.P.** disclosed no relevant relationships. **G.M.N.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: author disclosed personal fees from Bayer Health Care and Bracco and other from Three Palm Software. Other relationships: disclosed no relevant relationships. **H.S.M.** disclosed no relevant relationships. **L.J.E.** disclosed no relevant relationships. **M.D.S.** disclosed no relevant relationships.

References

1. Croshaw R, Shapiro-Wright H, Svensson E, Erb K, Julian T. Accuracy of clinical examination, digital mammogram, ultrasound, and MRI in determining postneoadjuvant pathologic tumor response in operable breast cancer patients. Ann Surg Oncol 2011;18(11):3160–3163.
2. Londero V, Bazzocchi M, Del Frate C, et al. Locally advanced breast cancer: comparison of mammography, sonography and MR imaging in evaluation of residual disease in women receiving neoadjuvant chemotherapy. Eur Radiol 2004;14(8):1371–1379.
3. Shin HJ, Kim HH, Ahn JH, et al. Comparison of mammography, sonography, MRI and clinical examination in patients with locally advanced or inflammatory breast cancer who underwent neoadjuvant chemotherapy. Br J Radiol 2011;84(1003):612–620.
4. Yeh E, Slanetz P, Kopans DB, et al. Prospective comparison of mammography, sonography, and MRI in patients undergoing neoadjuvant chemotherapy for palpable breast cancer. AJR Am J Roentgenol 2005;184(3):868–877.
5. Akazawa K, Tamaki Y, Taguchi T, et al. Preoperative evaluation of residual tumor extent by three-dimensional magnetic resonance imaging in breast cancer patients treated with neoadjuvant chemotherapy. Breast J 2006;12(2):130–137.
6. Montemurro F, Martincich L, De Rosa G, et al. Dynamic contrast-enhanced MRI and sonography in patients receiving primary chemotherapy for breast cancer. Eur Radiol 2005;15(6):1224–1233.
7. Weatherall PT, Evans GF, Metzger GJ, Saborrian MH, Leitch AM. MRI vs. histologic measurement of breast cancer following chemotherapy: comparison with x-ray mammography and palpation. J Magn Reson Imaging 2001;13(6):868–875.
8. Rosen EL, Blackwell KL, Baker JA, et al. Accuracy of MRI in the detection of residual breast cancer after neoadjuvant chemotherapy. AJR Am J Roentgenol 2003;181(5):1275–1282.
9. Ko ES, Han BK, Kim RB, et al. Analysis of factors that influence the accuracy of magnetic resonance imaging for predicting response after neoadjuvant chemotherapy in locally advanced breast cancer. Ann Surg Oncol 2013;20(8):2562–2568.
10. Lobbes MB, Prevos R, Smidt M, et al. The role of magnetic resonance imaging in assessing residual disease and pathologic complete response in breast cancer patients receiving neoadjuvant chemotherapy: a systematic review. Insights Imaging 2013;4(2):163–175.

11. Partridge SC, Gibbs JE, Lu Y, et al. MRI measurements of breast tumor volume predict response to neoadjuvant chemotherapy and recurrence-free survival. AJR Am J Roentgenol 2005;184(6):1774–1781.

12. Lobbes MB. Treatment response evaluation by MRI in breast cancer patients receiving neoadjuvant chemotherapy: there is more than just pathologic complete response prediction. Breast Cancer Res Treat 2012;136(1): 313–314.

13. Marinovich ML, Houssami N, Macaskill P, et al. Meta-analysis of magnetic resonance imaging in detecting residual breast cancer after neoadjuvant therapy. J Natl Cancer Inst 2013;105(5):321–333.

14. Esserman LJ, Berry DA, Cheang MC, et al. Chemotherapy response and recurrence-free survival in neoadjuvant breast cancer depends on biomarker profiles: results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657). Breast Cancer Res Treat 2012;132(3):1049–1062.

15. Hylton NM, Blume JD, Bernreuter WK, 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(3):663–672.

16. Hylton NM. Vascularity assessment of breast lesions with gadolinium-enhanced MR imaging. Magn Reson Imaging Clin N Am 1999; 7(2):411–420, x.

17. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45(2):228–247.

18. Esserman LJ, Berry DA, DeMichele A, et al. Pathologic complete response predicts recurrence-free survival more effectively by cancer subset: results from the I-SPY 1 TRIAL--CALGB 150007/150012, ACRIN 6657. J Clin Oncol 2012;30(26):3242–3249.

19. Symmans WF, Peintinger F, Hatzis C, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol 2007;25(28):4414–4422.

20. Hudis CA, Barlow WE, Costantino JP, et al. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. J Clin Oncol 2007; 25(15):2127–2132.

21. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med 2011;30(10):1105–1117.

22. Hosmer DW Jr, Lemeshow S, May S. Applied survival analysis: regression modeling of time to event data. 2nd ed. New York, NY: Wiley, 2008.

23. Wasser K, Sinn HP, Fink C, et al. Accuracy of tumor size measurement in breast cancer using MRI is influenced by histological regression induced by neoadjuvant chemotherapy. Eur Radiol 2003;13(6):1213–1223.

24. Denis F, Desbiez-Bourcier AV, Chapiron C, Arbion F, Body G, Brunereau L. Contrast enhanced magnetic resonance imaging underestimates residual disease following neoadjuvant docetaxel based chemotherapy for breast cancer. Eur J Surg Oncol 2004;30(10):1069–1076.

25. Warren RM, Bobrow LG, Earl HM, et al. Can breast MRI help in the management of women with breast cancer treated by neoadjuvant chemotherapy? Br J Cancer 2004;90(7):1349–1360.

26. King V, Goldfarb SB, Brooks JD, et al. Effect of aromatase inhibitors on background parenchymal enhancement and amount of fibroglandular tissue at breast MR imaging. Radiology 2012;264(3):670–678.

27. King V, Kaplan J, Pike MC, et al. Impact of tamoxifen on amount of fibroglandular tissue, background parenchymal enhancement, and cysts on breast magnetic resonance imaging. Breast J 2012;18(6):527–534.

28. Oksa S, Parkkola R, Luukkaala T, Mäenpää J. Breast magnetic resonance imaging findings in women treated with toremifene for premenstrual mastalgia. Acta Radiol 2009;50(9):984–989.

29. Carey LA, Dees EC, Sawyer L, et al. The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. Clin Cancer Res 2007;13(8):2329–2334.

30. Prowell T. Guidance to Industry. Pathologic Complete Response in Neoadjuvant Treatment of High-Risk Early-Stage Breast Cancer: Use as an Endpoint to Support Accelerated Approval. 2012.