

Breast DCE-MRI Kinetic Heterogeneity Tumor Markers: Preliminary Associations With Neoadjuvant Chemotherapy Response¹

Ahmed Ashraf^{*}, Bilwaj Gaonkar^{*}, Carolyn Mies[†], Angela DeMichele[‡], Mark Rosen^{*}, Christos Davatzikos^{*} and Despina Kontos^{*}

^{*}Department of Radiology, University of Pennsylvania School of Medicine, Philadelphia, PA, USA; [†]Department of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA, USA; [‡]Department of Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

Abstract

The ability to predict response to neoadjuvant chemotherapy for women diagnosed with breast cancer, either before or early on in treatment, is critical to judicious patient selection and tailoring the treatment regimen. In this paper, we investigate the role of contrast agent kinetic heterogeneity features derived from breast dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) for predicting treatment response. We propose a set of kinetic statistic descriptors and present preliminary results showing the discriminatory capacity of the proposed descriptors for predicting complete and non-complete responders as assessed from pre-treatment imaging exams. The study population consisted of 15 participants: 8 complete responders and 7 non-complete responders. Using the proposed kinetic features, we trained a leave-one-out logistic regression classifier that performs with an area under the receiver operating characteristic (ROC) curve (AUC) of 0.84 under the ROC. We compare the predictive value of our features against commonly used MRI features including kinetics of the characteristic kinetic curve (CKC), maximum peak enhancement (MPE), hotspot signal enhancement ratio (SER), and longest tumor diameter that give lower AUCs of 0.71, 0.66, 0.64, and 0.54, respectively. Our proposed kinetic statistics thus outperform the conventional kinetic descriptors as well as the classifier using a combination of all the conventional descriptors (i.e., CKC, MPE, SER, and longest diameter), which gives an AUC of 0.74. These findings suggest that heterogeneity-based DCE-MRI kinetic statistics could serve as potential imaging biomarkers for tumor characterization and could be used to improve candidate patient selection even before the start of the neoadjuvant treatment.

Translational Oncology (2015) 8, 154–162

Introduction

The use of neoadjuvant chemotherapy (NAC) in women diagnosed with breast cancer is gaining considerable acceptance [1,2]. It has been reported that NAC gives high clinical response of up to 70% to 98% and can result in a pathologic complete response (pCR) in 3% to 34% of patients [3–6]. On the other hand, it has also been reported that 2% to 30% of patients may not benefit clinically or pathologically [7]. As a result, the ability to distinguish between highly responsive and non-responsive patients is of critical importance for making treatment choices. Particularly, non-responsive patients, if detected early on, can avoid unnecessary side effects and can be routed to alternative therapies such as hormonal treatment and/or radiation therapy [8,9]. Moreover, identifying

potential responders can help in planning a more focused and timely treatment regimen [5].

Address all correspondence to: Ahmed Ashraf, PhD, Department of Radiology, University of Pennsylvania School of Medicine, Philadelphia, PA, USA.
E-mail: ahmed.bilal@alumni.cmu.edu

¹This work was supported by a pilot grant from the University of Pennsylvania Center for Biomedical Image Computing and Analytics.

Received 3 January 2015; Revised 21 March 2015; Accepted 24 March 2015

© 2015 The Authors. Published by Elsevier Inc. on behalf of Neoplasia Press, Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1936-5233/15

<http://dx.doi.org/10.1016/j.tranon.2015.03.005>

Traditionally, therapy response is evaluated by morphologic, clinical, and histopathologic assessment [10]. Imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) are used to assess tumor response primarily on the basis of tumor size reduction [11–13]. A widely used criterion for therapy response assessment is the Response Evaluation Criterion in Solid Tumors (RECIST) [14], which uses the longest diameter of the tumor as a basis. However, changes in tumor size due to therapy tend to appear after at least a couple of neoadjuvant treatment cycles [1]. Therefore, response as assessed by a reduction in tumor size can cause considerable delay in giving the appropriate treatment to partially responsive and non-responsive patients. As a result, devising methods for predicting early therapy response has been an active area of research in recent years.

A review on the use of MRI for predicting response to NAC is given in Marinovich et al. [15]. Among the work on early prediction of therapy response, correlations between NAC response and dynamic contrast-enhanced MRI (DCE-MRI) kinetic parameters such as the forward and backward volume transfer coefficients (K^{trans} and k_{ep}) have been reported [16–18]. Kinetic parameters derived from multiparametric MRI using DCE-MRI and diffusion-weighted MRI have also been explored for assessing the response to NAC treatment [19]. However, the changes in kinetic parameters in the above studies have predictive value after either one or two cycles of neoadjuvant treatment. Moreover, the estimation of volume transfer coefficients such as K^{trans} and k_{ep} involves a series of assumptions and models derived from the pharmacokinetics of the contrast agent distribution [20]. This leads to different estimates of the parameters due to different underlying assumptions [21], making the estimates less robust.

Hylton et al. [22] have explored morphologic and kinetic features derived from MRI such as the measurement of tumor longest diameter, changes in tumor volume, and peak signal enhancement ratio (SER) in post-treatment MR images as compared to pre-treatment images. Their analysis uses MR images acquired before treatment as well as images obtained after first, second, and third cycles of NAC treatment and shows that tumor volume, when represented as a fraction of the baseline tumor volume as assessed from pre-treatment images, has the greatest relative capacity to monitor NAC response. Another recent method from Tateishi et al. [23] investigates the prediction of pathologic response using positron emission tomography (PET) and DCE-MRI parameters. They use both the baseline imaging and images acquired after two NAC cycles. The results of Tateishi et al. report that the sensitivity of PET CT/MR after two cycles of NAC is not acceptable, while the specificity is high. One of the earlier works by Rieber et al. [24] also suggested that MR could provide information regarding response to therapy following the first cycle. Esserman et al. assessed morphologic MR phenotypes of tumors that correlated with NAC response [25]. However, the detection of these phenotypes was based on subjective and visual interpretation. Park et al. reported that pre-treatment values of mean apparent diffusion coefficient of responders were lower than that of non-responders [26].

The majority of the abovementioned methods are able to predict the NAC response only after at least a couple of treatment cycles. Among the methods that predict response based on pre-treatment imaging, the approach of Esserman et al. [25] is based on subjectively assessed morphologic phenotypes, while the method of Park et al. [26] uses an aggregate measure such as the mean apparent diffusion

coefficient of the tumor. Breast tumors have been shown to be highly heterogeneous lesions [27,28], and tumor heterogeneity is emerging as one of the major challenges in prognostication and successful treatment [27]. Breast cancer heterogeneity manifests in almost all aspects of the disease including histopathologic, molecular, and functional characteristics [28]. Traditional MRI analysis have mainly focused on aggregate kinetic measures [29–32] or aggregate diffusion measures [25]. However, such aggregate measures fall short of capturing the intratumor heterogeneity that is a crucial aspect of tumor characterization [27].

In this paper, we propose to explore heterogeneities in the contrast uptake within a breast tumor to predict response to NAC based on pre-treatment imaging. Our motivation to take this approach is based on reports suggesting that aggressive tumors have heterogeneous contrast uptake distribution perhaps due to inhomogeneous vascularization and angiogenesis [33–35]. The conventional kinetic descriptors either focus on aggregate kinetic measures or subsample the tumor, making little use of the information revealed by heterogeneities in the kinetic behavior. For instance, Bhooshan et al. [36] derive a single characteristic kinetic curve (CKC) and compute only one set of kinetic features from that curve. Similarly, Loiselle et al. [37] employ the maximum peak enhancement (MPE) as a single metric for the entire tumor. However, it has been suggested that heterogeneities in the contrast agent uptake may reflect important phenotypic aspects of underlying molecular intratumor heterogeneity [38–41]. Here, we propose to tap into this potentially rich set of information for the task of predicting response to neoadjuvant chemotherapy.

Material and Methods

Study Population

The study population consisted of a subset of patients from the ACIN 6657 ISPY trial [33], imaged at our institution between 2002 and 2006. Local institutional review board (IRB) approval was obtained before the study, and signed informed consent was obtained from all patients before enrollment. All patients presented with biopsy-proven T2 to T3 stage tumors of ductal carcinoma. The patients underwent standard neoadjuvant chemotherapy, which at the time of the study consisted of four cycles of adriamycin/cytosine, followed by four cycles of taxotere. Core biopsy and serum samples were collected at comparable times to the MRI scans and obtained pre-treatment, between 24 and 96 hours after the start of treatment, between treatment regimens (optional), and pre-surgery. The patients were declared as complete responders or partial responders after evaluating pathologic response [33]. pCR was defined as no invasive tumor present in either breast or axillary lymph nodes. The subset population consisted of 15 patients: 8 complete responders and 7 non-complete responders.

Imaging Parameters

The imaging parameters for ISPY-1 trial have been described before elsewhere [22]. We reproduce the imaging details for completeness. MRI was performed with a 1.5-T imaging unit by using a dedicated breast radiofrequency coil. Before the start of imaging, an intravenous catheter was inserted into each patient; patients were imaged in the prone position. The MRI protocol included a localization acquisition and T2-weighted sequence, followed by a DCE series. For T2-weighted imaging, a fast spin-echo

sequence with fat suppression was used in the sagittal orientation over the symptomatic breast only (two-dimensional spin echo; field of view, 1620 cm; section thickness, 3 mm; fat saturation; echo train length, 8 to 16; one echo; effective echo time, 80–140 milliseconds; repetition time, 4000–6000 milliseconds). For the contrast-enhanced series, high-spatial resolution (in-plane spatial resolution ≤ 1 mm) three-dimensional fat-suppressed T1-weighted imaging of the symptomatic breast was performed by using a gradient-echo sequence with a repetition time of 20 milliseconds or less, an echo time of 4.5 milliseconds, a flip angle of 45° or less, a 16- to 18-cm field of view, a minimum matrix of $256 \times 3 \times 192$, and 64 sections with a thickness of 2.5 mm or less. Imaging time for the T1-weighted sequence ranged from 90 to 100 seconds, with one data set acquired before injection of a gadolinium-based contrast agent and repeated at least two times immediately after injection. The contrast agent was administered at a dose of 0.1 mmol/kg body weight over 15 seconds, followed by a 10-ml saline flush over 15 seconds. The resulting temporal sampling of the center of k-space for the first contrast-enhanced phase was between 2 minutes 15 seconds and 2 minutes 30 seconds, providing image contrast most representative of this time point. An interphase delay between the first and the second contrast-enhanced phase was used as needed to result in temporal sampling of the second contrast-enhanced phase between 7 minutes 15 seconds and 7 minutes 45 seconds.

DCE-MRI Kinetic Features

For each patient, we selected the DCE-MRI exam done before the commencement of the treatment. Lesions were segmented in a semi-automated way by seeding an active contour snake [42]. Typically, DCE-MRI acquisition includes a pre-contrast image (captured before the injection of a contrast agent) and a number of

post-contrast images, captured at different time points after the injection of the contrast agent [43]. The uptake of the contrast agent by different tissues manifests itself in the form of contrast enhancement in post-contrast MR images, and the enhancement patterns have been suggested to be useful for tumor characterization [36,43].

A common way to quantify the enhancement pattern is to compute the relative enhancement as compared to the pre-contrast image [44,43]. By computing the relative enhancement on a pixel-by-pixel basis, we can construct pixel-wise maps of the relative contrast enhancement as follows:

$$\mathcal{E}(u, v, t) = \frac{I(u, v, t) - I(u, v, t_0)}{I(u, v, t_0)} \quad (1)$$

where $I(u, v, t)$ represents the intensity of pixel (u, v) captured at time t and t_0 is the pre-contrast time instant. For a particular pixel, the relative enhancement plotted as a function of time is defined as the kinetic curve [43]. In the literature (e.g., [44,43]), a number of basic features can be computed from this kinetic curve as illustrated in Figure 1A. On the basis of these features, we can derive a rich kinetic feature set by computing the pixel-wise map for each feature as follows:

- Peak enhancement (\mathcal{PE}):

$$\mathcal{PE}(u, v) = \max_t \mathcal{E}(u, v, t) \quad (2)$$

which represents the peak relative enhancement for every pixel as computed over all post-contrast time points.

- Time to peak (\mathcal{TTP}):

$$\mathcal{TTP}(u, v) = \arg \max_t \mathcal{E}(u, v, t) \quad (3)$$

which represents the time at which peak enhancement is achieved.

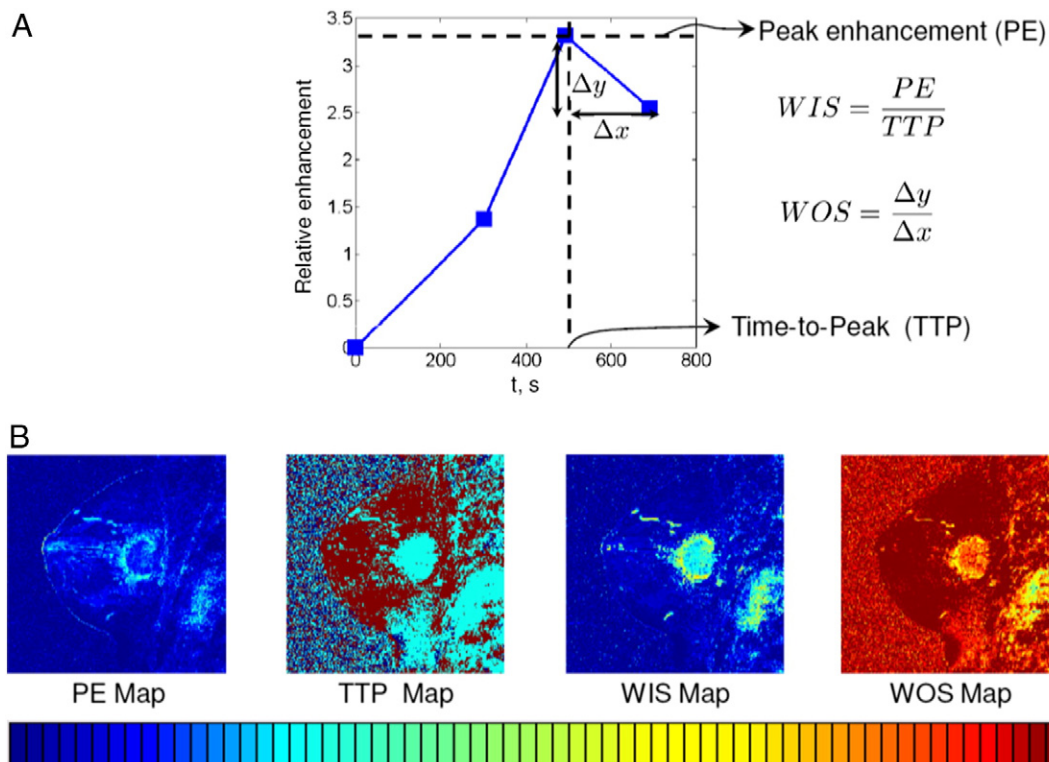


Figure 1. (A) Illustration of basic kinetic features for a single pixel. (B) Pixel-wise maps.

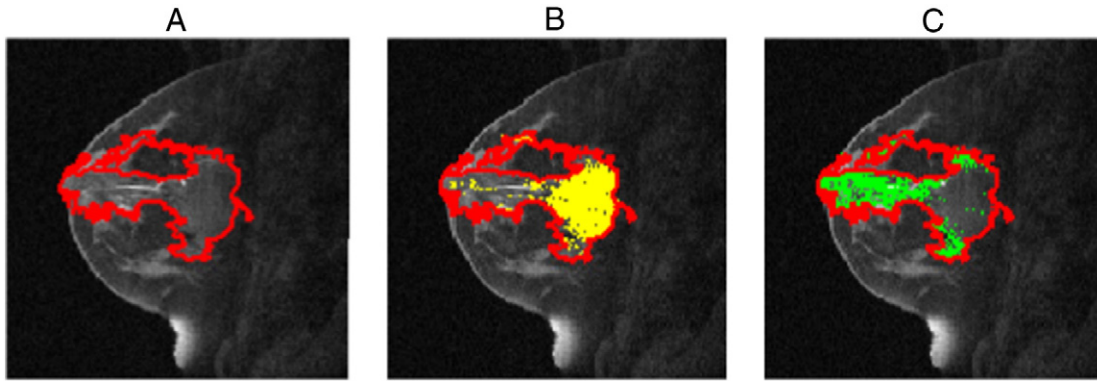


Figure 2. Illustration of kinetic pixel partitioning for two post-contrast time points. (A) Segmented lesion, (B) set 1 pixels that achieve peak enhancement at the first post-contrast time point (highlighted in yellow), and (C) set 2 pixels that achieve peak enhancement at the second post-contrast time point (highlighted in green).

- Wash-in slope (*WIS*):

$$WIS(u, v) = \begin{cases} \frac{\mathcal{PE}(u, v)}{\mathcal{TTP}(u, v) - t_0}; & \text{if } \mathcal{TTP}(u, v) \neq t_0 \\ 0; & \text{otherwise} \end{cases} \quad (4)$$

which is a measure of the initial uptake rate of the contrast agent for every pixel.

- Wash-out slope (*WOS*):

$$WOS(u, v) = \begin{cases} \frac{\mathcal{E}(u, v, t_M) - \mathcal{PE}(u, v)}{t_M - \mathcal{TTP}(u, v)}; & \text{if } t_M \neq \mathcal{TTP}(u, v) \\ 0; & \text{otherwise} \end{cases} \quad (5)$$

where t_M is the last post-contrast time instant. The wash-out slope captures the drop in the uptake rate of the contrast agent after the peak enhancement is achieved.

Sample pixel-wise maps of above features are depicted in Figure 1B.

Kinetic Statistic Features

To capitalize fully on the richness of the abovementioned pixel-wise kinetic feature maps, we propose to explore the contrast agent uptake heterogeneities within each of these feature maps. We draw our motivation from the fact that aggressive tumors tend to have heterogeneous contrast uptake distribution due to heterogeneous vascularization and angiogenesis [33,34]. One way to tap into the kinetic heterogeneity of a tumor is to divide the tumor pixels into clusters of homogeneity and look for patterns in feature statistics within every cluster. Specifically, we propose to partition the tumor pixels according to their kinetic behavior and compute statistics of kinetic feature maps within these partitions.

We extract these kinetic statistics in a two-phase process: In the first phase, the time to peak (TTP) for every pixel within the tumor is computed. We then cluster the pixels based on their TTP values. This

step partitions the pixels into as many sets as the number of post-contrast time points, i.e., set i consists of pixels that achieve their peak enhancement at the i th post contrast time point. Figure 2 illustrates these partitions for a sample case.

In the second step, pixel-wise maps of the following features are computed: peak enhancement (PE), wash-in slope (WIS), and wash-out-slope (WOS) (Figure 1). On the basis of these pixel groupings, we compute partition-wise kinetic statistics as follows: Let \mathcal{M} be the pixel partitioning such that \mathcal{M}_k represents the membership mapping of pixel k to its respective set. Given \mathcal{M} , one can derive the following set-wise statistics:

- Posterior probability of observing set i given the partition \mathcal{M} :

$$\mathcal{P}(\text{Set} = i | \mathcal{M}) = \frac{1}{N} \sum_{k=1}^N \delta(\mathcal{M}_k = i). \quad (6)$$

where $\delta(\mathcal{M}_k = i)$ is an indicator function that equals 1 when $\mathcal{M}_k = i$, and zero otherwise. N is the total number of pixels. These N pixels may come from an arbitrarily shaped segmentation mask specifying the lesion.

- Mean value of feature map j for set i :

$$\mu(i, j) = \frac{\sum_{k=1}^N f_j(k) \cdot \delta(\mathcal{M}_k = i)}{\sum_{k=1}^N \delta(\mathcal{M}_k = i)}. \quad (7)$$

where $f_j(k)$ is the value of the j -th feature map for k -th pixel, and the feature map can be any of those shown in Figure 1B.

- Variance of feature map j for set i :

$$\sigma^2(i, j) = \frac{\sum_{k=1}^N (f_j(k) - \mu(i, j))^2 \cdot \delta(\mathcal{M}_k = i)}{\sum_{k=1}^N \delta(\mathcal{M}_k = i)}. \quad (8)$$

On the basis of the above definitions, m pixel partitions and n feature maps would result into a total of $m(2n + 1)$ features. We aim to investigate the utility of these partition-wise kinetic statistics for the task of predicting response to neoadjuvant chemotherapy for breast cancer.

Table 1. Performance of Single Feature Classifiers. Column 2 Shows Feature Names Following the Notation of Equations (6) to (8), e.g., $\mu(1, PE)$ Represents the Mean Feature Map Value of Peak Enhancement for Set 1 Pixels. Column 3 Shows the AUC under the ROC for Each Feature

No.	Feature	ROC-AUC
1	$P(\text{Set} = 1 M)$	0.76
2	$P(\text{Set} = 2 M)$	0.76
3	$\mu(1, PE)$	0.77
4	$\mu(2, PE)$	0.81
5	$\sigma^2(1, PE)$	0.78
6	$\sigma^2(2, PE)$	0.73
7	$\mu(1, WIS)$	0.77
8	$\mu(2, WIS)$	0.81
9	$\sigma^2(1, WIS)$	0.78
10	$\sigma^2(2, WIS)$	0.73
11	$\mu(1, WOS)$	0.77
12	$\mu(2, WOS)$	0.53
13	$\sigma^2(1, WOS)$	0.75
14	$\sigma^2(2, WOS)$	0.54
15	Ellipticity	0.52
16	Circularity	0.51
17	Perimeter	0.53
18	Area	0.51

For feature extraction, the most representative slice in an image sequence was chosen as the slice for which the lesion area was maximized. The selection of the representative slice was confirmed by a board-certified radiologist. Given the representative slice, we computed the kinetic statistics as defined in the Results section. For this data set, we had two post-contrast time points, and on the basis of the TTP values, the pixels were partitioned into two sets ($m = 2$). Within these partitions, we computed statistics for three feature maps ($n = 3$), i.e., PE, WIS, and WOS. This resulted in $m(2n + 1) = 14$ kinetic statistics. As DCE-MRI morphologic features (e.g., tumor size and shape) have also been shown to be associated with different tumor grade [40], in our feature vector we also included tumor area and perimeter to measure tumor size and ellipticity and convexity to capture the shape of the tumor [45,46]. Therefore, in total we had 14 kinetic features and 4 morphologic features.

Classification Experiments

To evaluate the performance of individual features, we trained a leave-one-out single feature logistic regression classifier using each of the features individually. Receiver operating characteristic (ROC) analysis was performed, and the area under the ROC curve (AUC) was computed. To evaluate the collective performance of our features, we also trained a multivariate logistic regression classifier with leave-one-out cross-validation. For each leave-one-out loop, we performed sequential forward feature selection using the training examples for the current loop, i.e., all examples but the currently left out case. Using the selected features, a logistic regression classifier was learned on the basis of the training set and was tested on the unseen left out test example. The process was then repeated for all leave-one-out loops. ROC analysis was done to assess the performance of multivariate classifiers. To assess the importance of features, selection frequencies of features over all cross-validation folds were evaluated.

Comparison with Previous Work

We compare the predictive value of our features against the following commonly used MRI features:

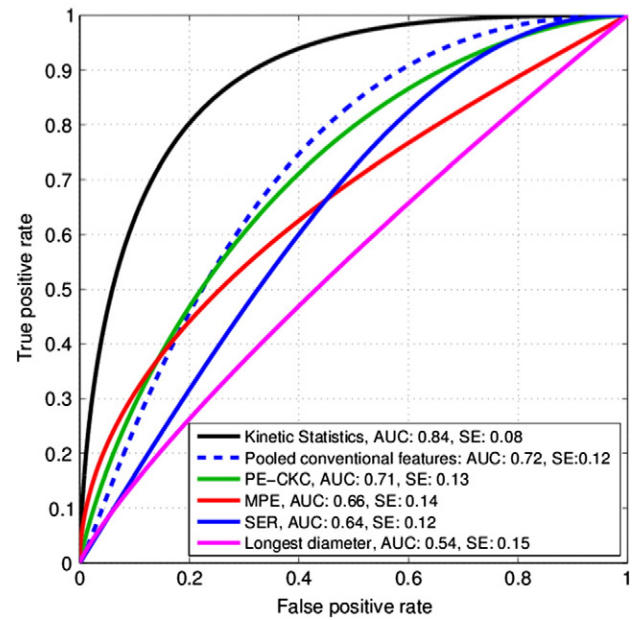


Figure 3. ROC comparison with standard features. The classifier based on our proposed kinetic statistics outperforms standard kinetic descriptors and the longest diameter that form the basis of the RECIST criterion. The dotted blue ROC represents the case when all the conventional features were pooled together.

- Kinetic features of the CKC [36,47]: On the basis of this method, kinetic curves were computed for all the tumor pixels, and each kinetic curve was represented as an M dimensional vector, where M is the number of time points. Fuzzy C means clustering was then run on these M dimensional vectors to group similar kinetic curves, and the centroid of each cluster was termed as a prototype kinetic curve. The number of centroids was selected using the same approach as Chang et al. [47], in which the authors set the number of centroids to either $N/160$ or 2, whichever is greater, where N is the number of tumor pixels. From among the prototype kinetic curves, the curve with the maximum initial uptake was chosen as the CKC, and the following kinetic features for the CKC were computed: peak enhancement, time to peak, wash-in slope, and wash-out slope. A leave-one-out logistic regression classifier similar to the one for multifeature classification scheme as described above was trained.
- Hotspot SER [43]: SER at every pixel was computed as $SER = \frac{S_1 - S_0}{S_2 - S_0}$, where S_0 , S_1 , and S_2 are the signal intensities at the pre-contrast, first post-contrast, and second post-contrast time points. A hotspot SER value was defined as the highest eight-connected voxel average over the tumor volume, and a leave-one-out logistic regression classifier was trained using hotspot SER values.
- MPE [37]: Kinetic curves for all the pixels were first computed, and the peak enhancement for each curve was evaluated. The MPE value for each lesion was then detected and a leave-one-out logistic regression classifier based on MPE was trained.
- Longest diameter [14]: The longest diameter of a tumor is a standard metric that forms the basis of the RECIST [14]. A leave-one-out logistic regression classifier was trained on the basis of the longest diameter values.
- Pooled conventional features: Finally, we trained a classifier model with leave-one-out feature selection by pooling all the conventional features described in this section (i.e., CKC, SER, MPE, and longest diameter).

Table 2. Comparison of Statistical Significance, Standard Errors, and 95% Confidence Intervals for the AUCs Corresponding to Classification Models Based on Proposed Features and Previously Used Standard Features

Features	AUC	Standard Error	P Value	95% Confidence Interval	
				Lower Bound	Upper Bound
Kinetic statistics	0.84	0.088	.015	0.697	1.000
Pooled conventional features	0.72	0.125	.154	0.455	0.994
PE-CKC	0.71	0.132	.165	0.436	0.992
MPE	0.66	0.144	.355	0.332	0.954
SER	0.64	0.122	.563	0.271	0.908
Longest diameter	0.54	0.152	.728	0.243	0.864

Results

AUC under ROCs for single feature classifiers are shown in Table 1. Single feature classifiers for 12 of 14 kinetic statistics perform with AUCs ranging from 0.73 to 0.81, while two of the kinetic features and all the morphologic features perform close to chance. The best single feature performance (AUC = 0.81) was given by the mean peak enhancement for set 2 pixels, $\mu(2, PE)$. This analysis suggests that partition-based kinetic statistics potentially possess univariate discriminatory power to predict responders versus non-complete responders from their pre-treatment images.

The ROC curve for multivariate classifier based on the proposed features along with a comparison against previously used features is shown in Figure 3. Classifier based on the proposed features has an AUC of 0.84 (SE = 0.088, P = .015) under the ROC curve. Comparison of statistical significance, standard errors, and 95% confidence intervals for the AUCs corresponding to classification models based on proposed features and previously used standard features is given in Table 2. The comparison against commonly used approaches shows that the classifier based on our proposed kinetic statistics outperforms classifiers using conventional descriptors as well as the classifier based on pooling all the conventional descriptors.

Selection frequencies of features over all cross-validation folds are given in Table 3. Most frequently selected features included mean peak enhancement for set 2 pixels, variance in peak enhancement for set 1 pixels, and mean wash-out slope for set 1 pixels. Representative contrast uptake patterns for the selected features along with their corresponding box plots are shown in Figure 4. A visualization matrix of squared correlation coefficients (R^2) between features is shown in Figure 5.

Discussion

In this paper, we have presented preliminary results showing that heterogeneity kinetic statistics derived from DCE-MRI feature maps (peak enhancement, wash-in slope, and wash-out slope) are significant predictors of response to neoadjuvant chemotherapy. We have shown that these statistics have predictive ability when derived from a DCE-MRI exam done before the commencement of the treatment. Most existing methods [16,1,48] become predictive after at least two cycles of neoadjuvant treatment. The ability to predict response before the start of the treatment can help in selecting the patients that are expected to benefit the most from the treatment, while routing the anticipated non-responders and partial responders to alternative therapies. Moreover, the classifier based on our proposed kinetic statistics outperforms the classifiers using standard kinetic descriptors. Our proposed statistics capitalize on the heterogeneities in the contrast agent uptake within the tumor and the improvement is in line with studies suggesting the importance of

Table 3. Selection Frequencies of Features over All Leave-One-Out Classification Loops. Column 2 Shows Feature Names Following the Notation of Equations (6) to (8), e.g., $\mu(1, PE)$ Represents the Mean Feature Map Value of Peak Enhancement for Set 1 Pixels. Column 3 Shows the Selection Frequency of Each Feature. Frequencies $\geq 50\%$ Are Shown in Bold

No.	Feature	Selection Frequency (%)
1	$P(\text{Set} = 1 M)$	0/15 (0.0%)
2	$P(\text{Set} = 2 M)$	0/15 (0.0%)
3	$\mu(1, PE)$	2/15 (13.3%)
4	$\mu(2, PE)$	14/15 (93.3%)
5	$\sigma^2(1, PE)$	12/15 (80.0%)
6	$\sigma^2(2, PE)$	1/15 (6.7%)
7	$\mu(1, WIS)$	0/15 (0.0%)
8	$\mu(2, WIS)$	0/15 (0.0%)
9	$\sigma^2(1, WIS)$	0/15 (0.0%)
10	$\sigma^2(2, WIS)$	0/15 (0.0%)
11	$\mu(1, WOS)$	12/15 (80.0%)
12	$\mu(2, WOS)$	0/15 (0.0%)
13	$\sigma^2(1, WOS)$	0/15 (0.0%)
14	$\sigma^2(2, WOS)$	0/15 (0.0%)
15	Ellipticity	0/15 (0.0%)
16	Circularity	0/15 (0.0%)
17	Perimeter	0/15 (0.0%)
18	Area	0/15 (0.0%)

kinetic heterogeneities for tumor characterization [43], as they can be potential markers of the underlying molecular intratumor heterogeneity [38–41].

Our results also indicate that different contrast uptake patterns exist for completely responding and partially responding tumors, including more pronounced enhancement, rapid washout, and heterogeneous uptake for the responding tumors (Figure 4), suggesting a more aggressive tumor phenotype, potentially due to higher vascularity and/or more progressed neo-angiogenic properties. For instance, the visualization and box plot in Figure 4A show that responding tumors have higher values of mean peak enhancement for set 2 pixels. The second row of Figure 4 shows that responding tumors have higher variance in the peak enhancement of their rapidly enhancing regions (i.e., TTP = 1). In Figure 4C, we show how washout characteristics differ between complete responders and partial responders. Specifically, responding tumors have a more steeper washout slope for their rapidly enhancing regions. This analysis suggests that tumors with pronounced and heterogeneous enhancement as well as rapid washout are more likely to respond to neoadjuvant chemotherapy. These results are in agreement with current oncologic literature [33,34], indicating that more biologically aggressive tumors have better response to neoadjuvant treatment.

While the majority of kinetic statistics (12 of 14) had univariate discriminatory capacity to predict responders versus partial responders, we should note that only three were consistently selected by the feature selection step, leading to a sparse classification model. This is primarily due to correlations among the extracted kinetic features. On the basis of the visualization of squared correlation coefficients between features in Figure 5, it can be seen that there are strong correlations between the corresponding kinetic statistics (mean and variance) of peak enhancement and wash-in slope. Since wash-in slope is a scalar multiple of peak enhancement ($WIS = PE/TTP$), it is straightforward to show that their respective statistics are also scalar multiples of each other (i.e., $\mu(i, WIS) = \mu(i, PE)/TTP$ and $\sigma^2(i, WIS) = \sigma^2(i, PE)/TTP^2$), and normalizing the features would lead to perfect correlations between them as represented in Figure 5.

Compared to the kinetic features, morphologic descriptors demonstrated poor performance. We must note, however, that morphologic

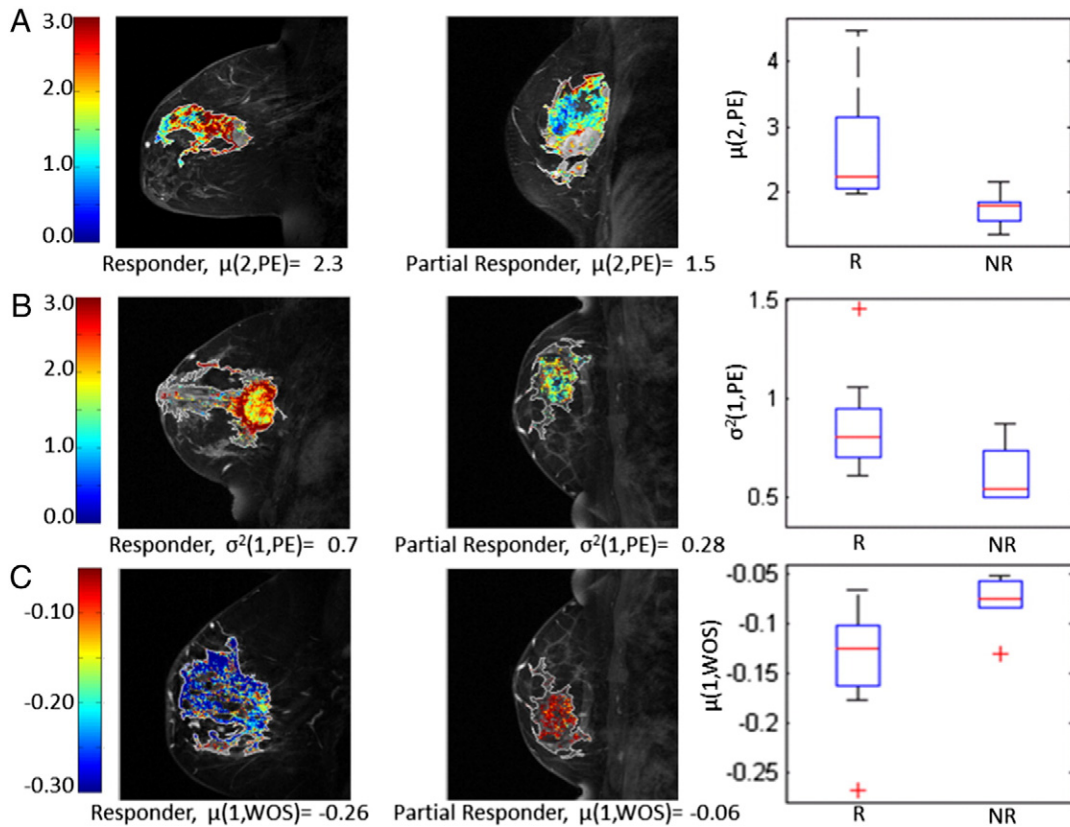


Figure 4. Representative contrast uptake patterns for the selected features. (A) Examples for pixel-wise maps of peak enhancement (PE) for set 2 pixels (TTP = second post-contrast time point). Complete responders have higher values of PE (i.e., red), causing the mean PE within set 2 pixels, $\mu(2, PE)$ to be higher than that of partial responders as depicted in the box plot for $\mu(2, PE)$. (B) Examples for pixel-wise maps of PE for set 1 pixels (TTP = first post-contrast time point). Complete responders have more variance in their set 1 PE values as depicted in the box plot for $\sigma^2(1, PE)$. (C) Examples for pixel-wise maps of wash-out slope (WOS) for set 1 pixels (TTP = first post-contrast time point). Complete responders have more rapid WOS (i.e., more negative shown as blue), compared to partial responders.

features have little or no correlation with kinetic features as illustrated by the visualization of the inter-feature correlations (Figure 5). This perhaps highlights the point that while their information might not be helpful for the particular task of pre-treatment therapy response

prediction, morphologic features capture different information compared to kinetic features. In this paper, we have focused on images captured before the initiation of the treatment begins. As the treatment progresses, changes in the morphology during the actual treatment process, rather than the *a priori* imaging, can represent a measure of response as has been explored in previous studies. Moreover, while RECIST longest diameter is traditionally used to evaluate the response to therapy, we also intended to investigate if tumor diameter before the initiation of the treatment could be a predictor of response as well, and an ROC-AUC of 0.5 suggests that it is not.

Important limitations of the current work must also be noted. First, although promising, our analysis has been done on a relatively small data set (15 patients), which has almost equal number of complete responders and partial responders, whereas in a standard NAC treatment regimen, the expected number of complete responders is around 25% [33], and therefore larger studies are warranted to confirm generalizability of our findings. Nevertheless, we have employed cross-validation in our experiments to assess generalization, and the selected features appear to be fairly consistent across different cross-validation loops. We should note that the temporal resolution of images used in this study is low with only three time points (i.e., one pre-contrast and two post-contrast). However, we would like to highlight that the proposed methodology is independent of the temporal resolution and the number of post-contrast time points. For future studies, images are being captured at higher temporal resolution. In addition, the features are

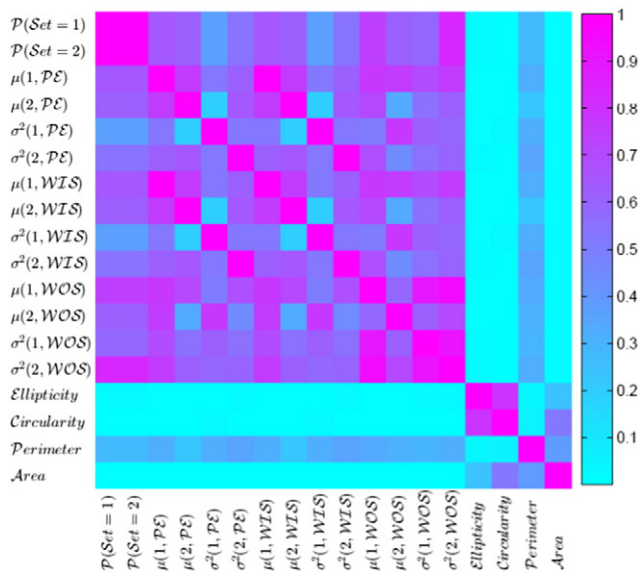


Figure 5. Visualization of the matrix of squared correlation coefficients (R^2) between features.

currently extracted from a representative two-dimensional slice of the primary lesion. Kinetic partitioning of the entire three-dimensional volume of the lesion could potentially lead to richer statistics, especially in the case of multifocal tumors, which may further improve the prediction of therapy response in the future. As mentioned earlier, pCR was assessed on the basis of no invasive cells in the excised tumor at the time of the surgery, which is the norm for other NAC studies. However, foci for ductal carcinoma in-situ (DCIS) might still be present, which should be investigated in future studies. Finally, integrating proposed imaging biomarkers with T2-weighted and diffusion-weighted imaging as well as with tumor histopathology and molecular subtypes holds the promise for better therapy response prediction and improved clinical decision making for individual patients and tumor subtypes.

Conclusions

In this paper, we have investigated the role of heterogeneity kinetic statistic features derived from breast DCE-MRI images for predicting neoadjuvant treatment response as assessed from imaging exams done before the treatment. We train a leave-one-out logistic regression classifier, which when tested on unseen test cases performs with AUC = 0.84 under the ROC curve. This analysis is based on a leave-one out feature selection and we show that kinetic statistics capturing the extent of contrast enhancement, uptake heterogeneity, and washout characteristics are among consistently selected features across cross-validation folds. We compare the predictive value of our features against that of commonly MR features including kinetics of CKC, hotspot SER, MPE, and longest diameter, which give lower AUCs of 0.71, 0.66, 0.64, and 0.54, respectively. These findings suggest that DCE-MRI heterogeneity kinetic statistics can be used to improve candidate patient selection even before the start of the neoadjuvant treatment. Larger studies are warranted to validate our findings in independent populations.

Acknowledgements

We thank the I-SPY-1 TRIAL team for the permission to use the I-SPY trial data and useful discussions on this research work.

References

- [1] Loo CE, Teertstra HJ, Rodenhuis S, van de Viver DA, Hannemann J, Muller DA, Peterers MJ, and Gilhuijs KG (2008). Dynamic contrast-enhanced MRI for prediction of breast cancer response to neoadjuvant chemotherapy: initial results. *AJR Am J Roentgenol* **191**, 1331–1338.
- [2] Le-Petross HC and Hylton N (2010). Role of breast MR imaging in neoadjuvant chemotherapy. *Magn Reson Imaging Clin N Am* **18**(2), 249–258.
- [3] Rastogi P, Anderson SJ, Bear HD, Geyer CE, Kahlenberg M, Robidoux A, Margolese R, Hoehn J, Vogel V, and Dakhil S, et al (2008). Preoperative chemotherapy: updates of National Surgical Adjuvant Breast and Bowel Project Protocols B-18 and B-27. *J Clin Oncol* **26**, 778–785.
- [4] Bonadonna G, Valagussa P, Brambilla C, Ferrari L, Moliterni A, Terenziani DA, and Zambetti M (1998). Primary chemotherapy in operable breast cancer: eight-year experience at the Milan Cancer Institute. *J Clin Oncol* **16**, 93–100.
- [5] Bear HD, Anderson S, Brown A, Smith R, Mamounas E, Fisher B, Margolese R, Theoret H, Soran A, and Wickerham D, et al (2003). The effect on tumor response of adding sequential preoperative docetaxel to preoperative doxorubicin and cyclophosphamide: preliminary results from National Surgical Adjuvant Breast and Bowel Project Protocol B-27. *J Clin Oncol* **21**, 4165–4174.
- [6] Makris A, Powles TJ, Ashley SE, Chang J, Hickish T, Tidy VA, Nash AG, and Ford HT (1998). A reduction in the requirements for mastectomy in a randomized trial of neoadjuvant chemoendocrine therapy in primary breast cancer. *Ann Oncol* **9**, 1179–1184.

- [7] Smith IC, Heys SD, Hutcheon AW, Miller I, Payne S, Gilbert F, Ah-See A, Eremin O, Walker L, and Sarkar T, et al (2002). Neoadjuvant chemotherapy in breast cancer: significantly enhanced response with docetaxel. *J Clin Oncol* **20**, 1456–1466.
- [8] Cuzick J, Sestak I, Pinder SE, Ellis IO, Forsyth S, Bundred NJ, Forbes JF, Bishop H, Fentiman IS, and George WD (2011). Effect of tamoxifen and radiotherapy in women with locally excised ductal carcinoma in situ: long-term results from the UK/ANZ DCIS trial. *Lancet Oncol* **21**(1), 21–29.
- [9] Correa C, McGale P, Taylor C, Wang Y, Clarke M, Davies C, Peto R, Bijker N, Solin L, and Darby S (2010). Overview of the randomized trials of radiotherapy in ductal carcinoma in situ of the breast. *J Natl Cancer Inst Monogr* **41**, 162–177.
- [10] Thoeny HC and Ross BD (2010). Predicting and monitoring cancer treatment response with diffusion-weighted MRI. *J Magn Reson Imaging* **32**(1), 2–16.
- [11] Chen JH, Bahri S, Mehta RS, Kuzucan A, Yu HJ, Carpenter PM, Feig SA, Lin M, Hsiang DJ, and Lane KT, et al (2011). Breast cancer: evaluation of response to neoadjuvant chemotherapy with 3.0-T MR imaging. *Radiology* **261**(3), 735–743.
- [12] McLaughlina R and Hylton N (2011). MRI in breast cancer therapy monitoring. *NMR Biomed* **24**, 712–720.
- [13] Arlinghaus LR, Li X, Levy M, Smith D, Welch EB, Gore JC, and Yankeelov TE (2010). Current and future trends in magnetic resonance imaging assessments of the response of breast tumors to neoadjuvant chemotherapy. *J Oncol*, 919620. <http://dx.doi.org/10.1155/2010/919620>.
- [14] Therasse P, Arbusck SG, Eisenhauer EA, Wanders DA, Kaplan RS, Rubinstein DA, Verweij J, Glabbeke DA, van Oosterom ATAT, and Christian MC, et al (2000). New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst* **92**(3), 205–216.
- [15] Marinovich ML, Sardanelli DA, Ciatto S, Mamounas DA, Brennan M, Macaskill DA, Irwig L, von Minckwitz DA, and Houssami N (2012). Early prediction of pathologic response to neoadjuvant therapy in breast cancer: systematic review of the accuracy of MRI. *Breast* **21**(5), 669–677.
- [16] Ah-See MLW, Makris A, Taylor NJ, Harrison M, Richman P, Burcombe R, Stirling J, d'Arcy JA, Collins DJ, and Pittam MR, et al (2008). Early changes in functional dynamic magnetic resonance imaging predict for pathologic response to neoadjuvant chemotherapy in primary breast cancer. *Clin Cancer Res* **14**(20), 6580–6589.
- [17] Li X, Arlinghaus L, Ayers GD, Chakravarthy AB, Abramson DA, Abramson VG, Atuegwu N, Farley J, Mayer IA, and Kelley MC, et al (2014). DCE-MRI analysis methods for predicting the response of breast cancer to neoadjuvant chemotherapy: pilot study findings. *Magn Reson Med* **71**(4), 1592–1602.
- [18] Etxano J, Insausti LP, Elizalde A, López-Vega DA, Plazaola A, and Martínez P (2014). Analysis of the changes induced by bevacizumab using a high temporal resolution DCE-MRI as prognostic factors for response to further neoadjuvant chemotherapy. *Acta Radiol* Oct 27. [pii: 0284185114556098, Epub ahead of print].
- [19] Li X, Abramson RG, Arlinghaus LR, Kang H, Chakravarthy AB, Abramson DA, Farley J, Mayer DA, Kelley MC, and Meszoely DA, et al (2015). Multiparametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer. *Invest Radiol* **50**(4), 195–204.
- [20] Cha S, Yang L, Johnson G, Lai A, Chen MH, Tihan T, Wendland M, and Dillon W (2006). Comparison of microvascular permeability measurements, K^{trans} , determined with conventional steady-state T1-weighted and first-pass T2*-weighted MR imaging methods in gliomas and meningiomas. *AJNR Am J Neuroradiol* **27**, 409–417.
- [21] Tofts PS, Brix G, Buckley DL, Evelhoch JL, Henderson E, Knopp MV, Larsson HB, Lee TY, N.A.N.A.M.Parker GJ, et al (1999). Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of diffusable tracer: standardized quantities and symbols. *J Magn Reson Imaging* **10**, 223–323.
- [22] Hylton NM, Blume JD, Bernreuter WK, Pisano DA, Rosen MA, Morris DA, Weatherall PT, Lehman CD, Newstead GM, and Polin S, et al (2012). Locally advanced breast cancer: MR imaging for prediction of response to neoadjuvant chemotherapy—results from ACRIN 6657/I-SPY TRIAL. *Radiology* **263**(3), 663–672.
- [23] Tateishi U, Miyake M, Nagaoka T, Terauchi T, Kubota K, Kinoshita T, Daisaki H, and Macapinlac DA (2012). Neoadjuvant chemotherapy in breast cancer: prediction of pathologic response with PET/CT and dynamic contrast-enhanced MR imaging—prospective assessment. *Radiology* **263**(1), 53–63.
- [24] Rieber A, Zeitler H, Rosenthal H, Gorich J, Kreienberg R, Brambs HJ, and Tomczak R (1997). MRI of breast cancer: influence of chemotherapy on sensitivity. *Br J Radiol* **70**(833), 452–458.
- [25] Esserman L, Kaplan E, Partridge S, Tripathy D, Rugo H, Park J, Hwang S, Kuerer H, Sudilovsky D, and Lu Y, et al (2001). MRI phenotype is associated

- with response to doxorubicin and cyclophosphamide neoadjuvant chemotherapy in stage III breast cancer. *Ann Surg Oncol* **8**(6), 549–559.
- [26] Park SH, Moon WK, Cho N, Song IC, Chang JM, Park IA, Han W, and Noh DY (2010). Diffusion-weighted MR imaging: pretreatment prediction of response to neoadjuvant chemotherapy in patients with breast cancer. *Radiology* **257**(1), 56–63.
- [27] Marusyk A and Polyak K (2010). Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta* **1805**(1), 105–117.
- [28] Almendro V and Fuster G (2011). Heterogeneity of breast cancer: etiology and clinical relevance. *Clin Transl Oncol* **13**, 767–773.
- [29] Chen W, Giger ML, Lan L, and Bick U (2004). Computerized interpretation of breast mri: investigation of enhancement-variance dynamics. *Med Phys* **31**, 1076–1082.
- [30] Lee SH, Kim JH, Cho N, Park JS, Yang Z, Jung YS, and Moon WK (2010). Multilevel analysis of spatiotemporal association features for differentiation of tumor enhancement patterns in breast DCE-MRI. *Med Phys* **37**.
- [31] Zheng Y, Englander S, Baloch S, Zacharaki EI, Fan Y, Schnall MD, and Shen D (2009). STEP: spatiotemporal enhancement pattern for MR-based breast tumor diagnosis. *Med Phys* **37**(7), 3192–3204.
- [32] Agner S, Soman S, Libfeld E, McDonald M, Thomas K, Englander S, Rosen MA, Chin D, Noshier J, and Madabhushi A (2011). Textural kinetics: a novel dynamic contrast-enhanced (DCE)-MRI feature for breast lesion classification. *J Digit Imaging* **24**(3), 446–463.
- [33] Esserman LJ, Berry DA, Cheang MCU, Yau C, Perou CM, Carey L, DeMichele A, Gray JW, Conway-Dorsey K, and Lenburg DA, et al (2012). 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* **132**(3), 1049–1062.
- [34] Loo CE, Straver ME, Rodenhuis S, Muller SH, Wesseling J, Peeters DA, and Gilhuijs KG (2011). Magnetic resonance imaging response monitoring of breast cancer during neoadjuvant chemotherapy: relevance of breast cancer subtype. *J Clin Oncol* **29**, 660–666.
- [35] Weinstein S and Rosen M (2010). Breast MR imaging: current indications and advanced imaging techniques. *Radiol Clin North Am* **48**(5), 1013–1042.
- [36] Bhooshan N, Giger ML, Jansen SA, Li H, Lan L, and Newstead GM (2010). Cancerous breast lesions on DCE MR images: computerized characterization for image based prognostic markers. *Radiology* **254**(3), 680–690.
- [37] Loiselle CR, Eby PR, Peacock S, Kim JN, and Lehman CD (2011). Dynamic contrast-enhanced magnetic resonance imaging and invasive breast cancer: primary lesion kinetics correlated with axillary lymph node extracapsular extension. *J Magn Reson Imaging* **33**(1), 96–101.
- [38] Szabó BK, Aspelin P, Kristoffersen MW, Tot T, and Boné B (2003). Invasive breast cancer: correlation of dynamic MR features with prognostic factors. *Eur Radiol* **13**(11), 2425–2435.
- [39] Tse GM, Chaiwun B, Wong KT, Yeung DK, Pang AL, Tang AP, and Cheung HS (2007). Magnetic resonance imaging of breast lesions—a pathologic correlation. *Breast Cancer Res Treat* **103**(1), 1–10.
- [40] Tozaki M (2004). Interpretation of breast MRI: correlation of kinetic and morphological parameters with pathological findings. *Magn Reson Med Sci* **3**(4), 189–197.
- [41] Ashraf AB, Daye D, Gavenonis S, Mies C, Feldman M, Rosen M, and Kontos D (2014). Identification of intrinsic imaging phenotypes for breast cancer tumors: Preliminary associations with gene expression profiles. *Radiology* **272**(2), 374–384.
- [42] Kass M, Witkin A, and Terzopoulos D (1988). Snakes: active contour models. *Int J Comput Vis* **1**(4), 321–331.
- [43] Hylton N (2006). MR imaging for assessment of breast cancer response to neoadjuvant chemotherapy. *Magn Reson Imaging Clin N Am* **14**(3), 383–389.
- [44] Lee SH, Kim JH, Cho N, Park JS, Yang Z, Jung YS, and Moon WK (2010). Multilevel analysis of spatiotemporal association features for differentiation of tumor enhancement patterns in breast DCE-MRI. *Med Phys* **37**, 3940.
- [45] Stojmenovic M and Nayak A (2007). Direct ellipse fitting and measuring based on shape boundaries. Proceedings of the 2nd Pacific Rim Conference on Advances in Image and Video Technology. , PSIVT'07Berlin/Heidelberg, Germany: Springer-Verlag; 2007. p. 221–235.
- [46] Chang RF, Wu WJ, Moon WK, and Chen DR (2005). Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors. *Breast Cancer Res Treat* **89**(2), 179–185.
- [47] Chen W, Giger ML, Bick U, and Newstead GM (2006). Automatic identification and classification of characteristic kinetic curves of breast lesions on DCE-MRI. *Med Phys* **33**, 2878–2887.
- [48] Partridge SC, Vanantwerp DA, Doot RK, Chai X, Kurland BF, Eby PR, Specht JM, Dunnwald LK, Schubert EK, and Lehman CD, et al (2010). Association between serial dynamic contrast-enhanced MRI and dynamic ¹⁸F-FDG PET measures in patients undergoing neoadjuvant chemotherapy for locally advanced breast cancer. *J Magn Reson Imaging* **32**(5), 1124–1131.