

# SCIENTIFIC REPORTS



OPEN

## A Radio-genomics Approach for Identifying High Risk Estrogen Receptor-positive Breast Cancers on DCE-MRI: Preliminary Results in Predicting OncotypeDX Risk Scores

Tao Wan<sup>1,2,3</sup>, B. Nicolas Bloch<sup>3</sup>, Donna Plecha<sup>4</sup>, Cheryl L. Thompson<sup>5</sup>, Hannah Gilmore<sup>6</sup>, Carl Jaffe<sup>3</sup>, Lyndsay Harris<sup>7</sup> & Anant Madabhushi<sup>2</sup>

To identify computer extracted imaging features for estrogen receptor (ER)-positive breast cancers on dynamic contrast enhanced (DCE)-MRI that are correlated with the low and high *OncotypeDX* risk categories. We collected 96 ER-positive breast lesions with low (<18,  $N=55$ ) and high (>30,  $N=41$ ) *OncotypeDX* recurrence scores. Each lesion was quantitatively characterized via 6 shape features, 3 pharmacokinetics, 4 enhancement kinetics, 4 intensity kinetics, 148 textural kinetics, 5 dynamic histogram of oriented gradient (DHoG), and 6 dynamic local binary pattern (DLBP) features. The extracted features were evaluated by a linear discriminant analysis (LDA) classifier in terms of their ability to distinguish low and high *OncotypeDX* risk categories. Classification performance was evaluated by area under the receiver operator characteristic curve (Az). The DHoG and DLBP achieved Az values of 0.84 and 0.80, respectively. The 6 top features identified via feature selection were subsequently combined with the LDA classifier to yield an Az of 0.87. The correlation analysis showed that DHoG ( $\rho=0.85$ ,  $P<0.001$ ) and DLBP ( $\rho=0.83$ ,  $P<0.01$ ) were significantly associated with the low and high risk classifications from the *OncotypeDX* assay. Our results indicated that computer extracted texture features of DCE-MRI were highly correlated with the high and low *OncotypeDX* risk categories for ER-positive cancers.

In the United States, approximately 70% of all breast cancer patients are diagnosed with estrogen receptor (ER)-positive primary tumors, and many of these women will receive hormonal therapy and adjuvant chemotherapy<sup>1,2</sup>. Identifying patients who can be treated with hormonal therapy alone would limit toxicity to those most likely to benefit<sup>3</sup>. Currently, the selection of individualized therapy for patients with ER-positive, human epidermal growth factor receptor type 2 (HER2)-negative early stage breast cancers is guided by the *OncotypeDX* gene expression assay (Genomic Health Inc., Redwood City, CA)<sup>4</sup>, which produces a recurrence score between 0–100 to predict the likelihood of disease recurrence with tamoxifen and improves the oncologist's ability to estimate benefit from adjuvant hormonal and chemotherapy<sup>5,6</sup>. However, the *OncotypeDX* assay is performed on tissue biopsy specimens and involves additional tumor handling and shipping of tissue, and delays in time to treatment, all of which add costs and anxiety for patients<sup>3,7</sup>.

Computer-aided diagnosis (CAD) systems have been previously used with dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) of the breast to predict the presence or absence of cancer<sup>8</sup>. These CAD tools aim to increase the accuracy of diagnosis of breast cancer on MRI while also aiming to reduce inter-observer variability. However, more recently a number of groups have begun to focus on identifying computer extracted

<sup>1</sup>School of Biological Science and Medical Engineering, Beihang University, Beijing, China. <sup>2</sup>Dept. of Biomedical Engineering, Case Western Reserve University, OH, USA. <sup>3</sup>Dept. of Radiology, Boston University School of Medicine, MA, USA. <sup>4</sup>Dept. of Radiology, University Hospitals Case Medical Center, OH, USA. <sup>5</sup>Dept. of Nutrition, Case Western Reserve University, OH, USA. <sup>6</sup>Division of Anatomic Pathology, Case Western Reserve University, OH, USA. <sup>7</sup>Division of Hematology and Oncology, Seidman Cancer Center, OH, USA. Correspondence and requests for materials should be addressed to T.W. (email: taowan@buaa.edu.cn)

Feature Class	Feature Name	Az	PPV	NPV	$\rho^*$	P value	RSD%	err
DHoG	4 bins	0.84 (0.74, 0.94)	0.81 (0.76, 0.86)	0.87 (0.81, 0.93)	0.85 (0.83, 0.87)	0.006	4.76	0.157
	6 bins	0.82 (0.71, 0.93)	0.78 (0.70, 0.86)	0.85 (0.78, 0.92)	0.82 (0.79, 0.85)	0.008	6.09	0.174
DLBP	256 bins	0.80 (0.70, 0.90)	0.74 (0.68, 0.80)	0.85 (0.80, 0.90)	0.83 (0.80, 0.86)	0.008	5.01	0.141
	128 bins	0.79 (0.67, 0.91)	0.74 (0.66, 0.82)	0.83 (0.76, 0.90)	0.83 (0.81, 0.85)	0.013	7.59	0.184
PK	$K^{trans}$	0.74 (0.60, 0.88)	0.70 (0.63, 0.77)	0.78 (0.70, 0.86)	0.79 (0.73, 0.85)	0.021	8.11	0.202
	$K_{ep}$	0.70 (0.55, 0.85)	0.71 (0.61, 0.81)	0.66 (0.61, 0.71)	0.71 (0.67, 0.75)	0.032	8.57	0.245
EK	Uptake rate	0.72 (0.59, 0.85)	0.63 (0.59, 0.67)	0.74 (0.67, 0.81)	0.65 (0.61, 0.69)	0.064	7.64	0.211
	Time to peak	0.63 (0.52, 0.74)	0.56 (0.51, 0.60)	0.65 (0.57, 0.73)	-0.52 (-0.47, -0.57)	0.212	7.94	0.298
TK	Haralick (Energy)	0.70 (0.57, 0.83)	0.64 (0.59, 0.69)	0.71 (0.63, 0.79)	0.73 (0.70, 0.76)	0.017	7.86	0.257
	Kirsch (Magnitude)	0.68 (0.52, 0.84)	0.60 (0.54, 0.66)	0.72 (0.65, 0.79)	0.72 (0.67, 0.77)	0.052	8.82	0.319
IK	1 <sup>st</sup> fitting coefficient	0.64 (0.52, 0.76)	0.60 (0.53, 0.67)	0.64 (0.58, 0.70)	-0.43 (-0.37, -0.49)	0.286	10.16	0.326
	4 <sup>th</sup> fitting coefficient	0.63 (0.52, 0.74)	0.58 (0.51, 0.65)	0.64 (0.54, 0.74)	-0.39 (-0.32, -0.46)	0.483	8.73	0.293
Shape	Compactness	0.64 (0.53, 0.75)	0.58 (0.51, 0.65)	0.66 (0.60, 0.72)	-0.57 (-0.52, -0.62)	0.338	7.81	0.334
	Normalized average radial distance ratio	0.60 (0.52, 0.72)	0.53 (0.48, 0.58)	0.67 (0.59, 0.75)	0.53 (0.48, 0.58)	0.502	8.33	0.377

**Table 1. The best two identified features in each feature class associated with their performance measures in distinguishing low and high risk estrogen receptor (ER)-positive breast cancers.** Note. -Numbers in parentheses are 95% confidence intervals. Az = area under the receiver operating characteristic curve; PPV = positive predictive value; NPV = negative predictive value; DHoG = dynamic histogram of oriented gradient; DLBP = dynamic local binary pattern; PK = pharmacokinetics; EK = enhancement kinetics; TK = textural kinetics; IK = intensity kinetics; RSD = relative standard deviation; *err* = error rate of classification.  $\rho^*$  denotes correlation coefficient.

Parameters	Site I		Site II		P value
	Low (<18)	High (>30)	Low (<18)	High (>30)	
OncotypeDX recurrence score					
No. of Patients (N = 96)	12 (12%)	5 (5%)	43 (45%)	36 (38%)	
Age (y) <sup>*</sup>	52 (37–68)	47 (36–55)	55 (40–77)	54 (29–70)	0.27
Lesion Size (mm) <sup>*</sup>	13 (12–30)	21 (7–33)	18 (5–50)	17 (9–40)	0.18
Patient ethnicity					
White	7 (7%)	2 (2%)	36 (38%)	29 (31%)	0.24
African American	3 (3%)	0	7 (7%)	7 (7%)	0.12
Unknown	2 (2%)	3(3%)	0	0	0.08
PR status					
Positive	12 (12%)	3 (3%)	42 (44%)	26 (27%)	0.43
Negative	0	2 (2%)	1 (1%)	10 (11%)	0.21
HER2 status					
Positive	7 (7%)	4 (4%)	18 (19%)	29 (31%)	0.48
Negative	5 (5%)	1 (1%)	25 (26%)	7 (7%)	0.37
Histologic Tumor Grade					
Low	4 (4%)	1 (1%)	10 (11%)	8 (8%)	0.15
Moderate	8 (8%)	2 (2%)	29 (31%)	21 (22%)	0.57
High	0	2 (2%)	4 (4%)	7 (7%)	0.19
Tumor type					
IDC	8 (8%)	3 (3%)	33 (35%)	22 (23%)	0.32
ILC	3 (3%)	0	6 (6%)	11 (12%)	0.16
Mixed	1 (1%)	2 (2%)	4 (4%)	3 (3%)	0.09

**Table 2. Characteristics of patients with estrogen receptor (ER)-positive breast cancers.** Note. -Unless otherwise indicated, data are numbers of patients, with percentages in parentheses. IDC = invasive ductal carcinoma; ILC = invasive lobular carcinoma. <sup>\*</sup>Data are means, with ranges in parentheses.

features on MRI that might be associated with underlying biology molecular subtype and risk of recurrence of the tumor<sup>9–11</sup>. A radiogenomic approach presented by Yamamoto *et al.*<sup>12</sup> found a significant correlation between breast MRI (1.5 Tesla) features and a number of important breast cancer related gene sets. Correlation studies by Vassiou *et al.*<sup>13</sup> and Chang *et al.*<sup>14</sup> showed that DCE-MRI (1.5 Tesla) based imaging features, such as tumor margin, enhancement pattern, and kinetic characteristics, were associated with pathological prognostic factors for the prediction of clinical outcome during treatment of breast cancer. A recent study conducted by Sutton *et al.*<sup>15</sup> showed that two MRI (1.5 or 3.0 Tesla) derived statistical image features were significantly correlated with the

median *OncotypeDX* recurrence scores with a range of 0–45. Ashraf *et al.*<sup>9</sup> presented a method for identifying correlation between computer-extracted morphologic and kinetic features from DCE-MRI sequences obtained at a 1.5T magnet with validated prognostic gene expression profiles of breast cancers. In addition, Agner *et al.*<sup>8</sup> presented an approach called textural kinetics (TK), which involved measuring dynamic changes in breast lesion texture during contrast uptake. These TK features were able to separate different molecular subtypes of breast cancers (triple negative, ER-positive, HER2-positive, fibroadenoma) on DCE-MRI obtained at either 1.5T or 3T<sup>16</sup>.

In this work, we investigate the ability of TK features on DCE-MRI to distinguish ER-positive breast cancers between low and high *OncotypeDX* risk categories (i.e., *OncotypeDX* recurrence score <18 and *OncotypeDX* recurrence score >30). Our approach is different from the work presented by Ashraf *et al.*<sup>9</sup>, in which dynamic features were computed based on the estimation of parameters on time-intensity curve, e.g., peak enhancement, wash-in and wash-out slope. Our approach is also different from the work published by Sutton *et al.*<sup>15</sup>, in which the image features (i.e., morphological, static first-order, and Haralick texture features) were extracted from breast lesions on pre- and three post-contrast MR images. In this study, we focus on TK features that allow for characterization of dynamic texture changes, specifically texture involving dynamic histograms in tumors on 1.5 Tesla DCE-MRI. Also our TK features are different from the approach by Agner *et al.*<sup>8,16</sup>, in that it involves quantification of kinetic texture in a new way - dynamic histogram of oriented gradients (DHoG) and dynamic local binary patterns (DLBP).

The histogram of oriented gradients (HoG)<sup>17</sup>, local binary patterns (LBP)<sup>18</sup>, and their spatio-temporal representations<sup>19,20</sup> are image texture features that have been previously employed for detecting breast masses on mammographic images<sup>21,22</sup>. Unlike the approach in Agner *et al.*<sup>8,16</sup> which attempted to fit a single parametric curve to explain the changes in lesion texture during the contrast wash-in and wash-out, the DHoG and DLBP approaches allow for construction of a unique lesion signature that captures the frequency of occurrence of different spatio-temporal textural patterns.

In this study, we evaluate the ability of the DHoG and DLBP features extracted from DCE-MRI at 1.5 Tesla, via a linear discriminant analysis (LDA) classifier to distinguish low and high risk ER-positive breast cancers, risk having been established via the *OncotypeDX* assay.

## Results

Table 1 shows the best two identified features in each feature class (shape, PK, EK, IK, TK, DHoG, DLBP) associated with Az, PPV, NPV,  $\rho$  (correlation coefficient), RSD, and *err* (error rate of classification). The numbers of bins used in the DHoG and DLBP features were {2, 4, 6, 8, 10} and {8, 16, 33, 64, 128, 256}, respectively. Higher  $\rho$  value indicates a stronger relationship between risk stratification via the features and *OncotypeDX*. The post-hoc power analysis revealed no significant difference between features identified from imaging data acquired at Sites I and II which were found to discriminate high versus low *OncotypeDX* risk categories. The heat map (Fig. 1) shows the values of all the features listed in Table 1 for all the patient studies. The values of best two identified features in each feature class and the six top performing features obtained from the LDA based feature selection method are listed in Supplementary Table S1 online.

**Shape Features for Discriminating ER-positive Breast Cancers.** The feature compactness used to measure the speculation of tumor margin yielded the best discriminability among the computerized shape features. The lower values of compactness for the high *OncotypeDX* breast cancers ( $-20.76 \pm 7.32$ ) compared with the low *OncotypeDX* cancers ( $-19.05 \pm 9.14$ ) appears to suggest that higher *OncotypeDX* score cancers tend to be associated with more speculation compared to cancers with low *OncotypeDX* risk scores. Table 1 appears to suggest that shape features are less useful for differentiating between low and high *OncotypeDX* risk score ER-positive lesions compared to pharmacokinetic and textural kinetic features (i.e., PK, EK, IK, TK, DHoG, and DLBP).

**PK Features for Discriminating ER-positive Breast Cancers.** Among the PK parameters,  $K^{trans}$  was found to be the most effective in distinguishing low and high *OncotypeDX* risk score ER-positive breast lesions (Fig. 2). While compared to  $K_{ep}$ ,  $K^{trans}$  appeared to have a stronger correlation with the *OncotypeDX* risk scores.  $K_{ep}$  appeared to be more predictive in identifying lesions on DCE-MRI that had a high *OncotypeDX* risk score.

**EK Features for Discriminating ER-positive Breast Cancers.** The EK features were extracted to quantitatively characterize the contrast enhancement patterns within ROIs in the lesion. Although the uptake rate achieved the best classification performance among all the EK features, these features generally had weaker correlation with the *OncotypeDX* risk scores compared to the TK features.

**IK and TK Features for Discriminating ER-positive Breast Cancers.** The TK features outperformed the IK features by up to 10% in terms of Az and NPV (Table 1). Further, the TK features had a higher degree of positive correlation with the *OncotypeDX* risk categories compared to the IK features which had a lower degree of negative correlation. These trends are consistent with the performance of the classifier (Table 1).

**DHoG and DLBP Features for Discriminating ER-positive Breast Cancers.** Both DHoG and DLBP yielded good discrimination for separating lesions corresponding to the low and high *OncotypeDX* risk categories, achieving the highest values of Az, PPV, and NPV among all computer extracted MRI features. The DHoG and DLBP features were significantly correlated with the *OncotypeDX* risk categories.

Figures 3 and 4 illustrate the textural kinetic curves corresponding to contrast uptake and enhancement appearance of the DHoG and DLBP features for representative low and high *OncotypeDX* risk breast lesions.

Feature Class	Lesion Feature	Definition
Shape ( $k = 6$ )	Area overlap ratio	Quantitative measures on lesion shape and lesion margin
	Variance of distance ratio, Compactness, Smoothness	
	Normalized average radial distance ratio	
	Standard deviation of normalized distance ratio	
PK ( $k = 3$ )	$K^{trans}$	Transfer constant between plasma and tissue compartments
	$V_e$	The extracellular extravascular volume fraction
	$K_{ep}$	The ratio of $K^{trans}/V_e$
EK ( $k = 4$ )	Maximal uptake, Time to peak	Transfer constant between plasma and tissue compartments
	Uptake rate, Washout rate	
IK ( $k = 4$ )	Third polynomial fitting on intensity curve	Intensity kinetic descriptors
TK - first order statistics ( $k = 48$ )	Mean, Median	Region intensity statistics derived from lesion area Window size, $w \in \{3, 5, 7\}$
	Range	
	Standard deviation	
TK - Sobel filter ( $k = 12$ )	x-direction gradient, y-direction gradient	Edge detectors
	Magnitude of gradient	Window size is $3 \times 3$
TK - Kirsch filter ( $k = 36$ )	Directions: $0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4$	Non-linear edge detector through eight compass directions
	Magnitude of the Kirsch operator	
TK - Haralick ( $k = 52$ )	Contrast energy, Contrast inverse moment	Features derived from grey-level co-occurrence matrices
	Contrast average, Contrast variance	
	Contrast entropy	
	Intensity average, intensity variance, intensity entropy	
	Entropy, Energy, Correlation	
	Information Measure 1, Information Measure 2	
DHoG ( $k = 5$ )	The number of bins: 2, 4, 6, 8, 10	Histogram based descriptor for gradient orientation on DCE-MRI
DLBP ( $k = 6$ )	The number of bins: 8, 16, 32, 64, 128, 256	Dynamic local binary pattern features based on texture spectrum

**Table 3. Description of all features used to distinguish low and high risk estrogen receptor (ER)-positive breast cancers.** PK = pharmacokinetics; EK = enhancement kinetics; IK = intensity kinetics; TK = textural kinetics; DHoG = dynamic histogram of oriented gradient; DLBP = dynamic local binary pattern. 'k' denotes the number of features.

The average features values of the top DHoG and DLBP features obtained across all patient studies are shown in Fig. 5(a,b) in comparison with the PK features (Fig. 5(c)).

**Performance of the Feature Combination for Discriminating ER-positive Breast Cancers.** The top performing features from each of the 4 classes, including PK ( $K^{trans}$ ), TK (Energy, Sobel x-direction gradient), DHoG (4-bin, 6-bin), and DLBP (256-bin), were combined using a linear discriminant classifier. The combined LDA classifier was able to differentiate between low and high OncotypeDX ER-positive breast cancers with an Az of 0.87 (95% confidence interval: 0.78, 0.96), PPV of 0.76 (95% confidence interval: 0.68, 0.84), NPV of 0.89 (95% confidence interval: 0.83, 0.95), RSD of 5.75, and *err* of 0.147.

**Stability and Predictive Performance.** The stability and error of LDA classification were measured by the relative standard deviation and the inverse power law model. It can be observed that DHoG and DLBP outperformed the other feature classes (i.e., shape, PK, EK, TK, IK) and achieved the smallest RSD values and error rates. The IK feature (1<sup>st</sup> fitting coefficient) has the highest RSD values indicating the lowest stability in classification. The shape features produced the largest error rates among all the feature classes.

## Discussion

We presented a computerized image analysis framework for identification of breast MR imaging markers to distinguish between low and high risk ER-positive breast cancers via a correlation of computer extracted DCE-MRI attributes and the OncotypeDX assay. Although tumor margin, tumor size, rim enhancement on DCE-MRI have been previously correlated with pathological factors and have been reported to be associated with disease outcome<sup>11-14,23</sup>, to the best of our knowledge, this is the first attempt to investigate the association between textural kinetic features on DCE-MRI with OncotypeDX recurrence scores for ER-positive breast cancers. This is important because the OncotypeDX is an assay with proven clinical utility that has been shown to be both prognostic and predictive in ER-positive breast cancers<sup>24</sup>. Hence by demonstrating computer extracted imaging features on DCE-MRI can predict the OncotypeDX risk category of the lesion, we might be able to non-invasively identify which patients would benefit from adjuvant therapy. This could pave the way for non-invasive risk assessment of the lesion even prior to biopsy. Furthermore, this is the first systematic comparison of various kinetic (PK, EK, IK,

TK, DHoG and DLBP) and shape features, to discriminate high and low *OncotypeDX* categories of ER-positive breast lesions.

There has been recent interest in identifying radiogenomic correlates of breast lesions on MRI. In<sup>16</sup>, Agner *et al.* showed that textural kinetic features extracted from routine clinical DCE-MRI appeared to be associated with the biologic heterogeneity and molecular subtype of breast cancers. Giger *et al.*<sup>25,26</sup> computed enhancement kinetic features, such as time to peak, uptake rate, maximal uptake, from a characteristic time course curve to distinguish benign and malignant breast masses. Ashraf *et al.*<sup>9</sup> also utilized curve-based kinetic features to construct breast DCE-MRI phenotypes and showed their correlation with the *OncotypeDX* assay.

In this study, we attempted to identify whether there was an association between textural kinetic features extracted from ER-positive breast lesions on 1.5 Tesla DCE-MRI and their corresponding *OncotypeDX* risk categories. While Ashraf *et al.*<sup>9</sup> focused on the association between lesion shape, contrast kinetic features and spatial heterogeneity features and the continuous *OncotypeDX* recurrence scores, our approach was focused on evaluating the ability of quantitative image features and spatio-temporal patterns within the lesion to distinguish between the low (<18 risk score) and high (>30 risk score) *OncotypeDX* risk categories. Additionally in conjunction with a LDA classifier, the textural kinetic features yielded an  $Az = 0.84$  in distinguishing low and high *OncotypeDX* risk category lesions, compared to Ashraf *et al.*<sup>9</sup> where the  $Az$  was 0.77.

Our approach was also different from that of Agner *et al.*<sup>8</sup> in that we employed two new textural kinetic features, DHoG and DLBP, which unlike EK, IK and TK features, seek to capture contextual textural changes during contrast uptake by considering changes in spatial intensity patterns within divided grid cells in the lesion ROI. Unlike the approach in Agner *et al.*<sup>16</sup> which attempted to fit a single parametric curve to characterize the temporal changes in lesion texture, the DHoG and DLBP approaches capture the frequency of occurrence of different spatio-temporal textural patterns within the lesion.

A systematic and quantitative analysis of different computer extracted features demonstrated that curve-based kinetic features (i.e., EK, TK, IK) were less discriminating compared to the other three feature classes (i.e., DHoG, DLBP, PK) in distinguishing high and low *OncotypeDX* risk score ER-positive cancers. Consistently, the feature combination identified through the feature selection process contained the important features from 4 feature classes (DHoG, DLBP, PK, and TK). While the PK features showed moderate correlation, lesion shape features were even less correlated with the *OncotypeDX* risk categories for the lesions evaluated. The DHoG and DLBP appeared to be the most discriminative features in differentiating low and high *OncotypeDX* risk score ER-positive breast lesions on DCE-MRI. Figures 3 and 4 which show the normalized mean DHoG and DLBP curves plotted as a function of contrast uptake, appear to illustrate a high degree of heterogeneity in high *OncotypeDX* risk score cancers compared to low *OncotypeDX* risk score cancers. The corresponding color-coded DHoG and DLBP feature maps at peak enhancement (Figs 3 and 4) also suggest that high *OncotypeDX* risk score breast cancers may appear to be more heterogeneous at peak contrast compared to low *OncotypeDX* risk score cancers. The Spearman's rank correlation test showed that DHoG and DLBP are significantly correlated (DHoG:  $\rho = 0.85$ ,  $P < 0.01$ ; DLBP:  $\rho = 0.83$ ,  $P < 0.01$ ) with the high and low *OncotypeDX* risk score categories. These results are consistent with the findings of Ashraf *et al.*<sup>9</sup>, who showed that DCE-MRI based heterogeneity kinetic features were correlated with *OncotypeDX* recurrence scores ( $\rho = 0.71$ ,  $P < 0.001$ ). However unlike Ashraf *et al.* where image data from only a single institute was considered, our approach included image data from two different clinical sites.

Our study did have its limitations, and as such, it is important to acknowledge that this is a preliminary study with need for additional independent validation of our initial findings. Additionally, we only included those patients having low (<18) and high (>30) *OncotypeDX* recurrence scores and excluded intermediate risk scores (>18 and <30) as the contrast was greatest between these categories and further work is needed to evaluate the intermediate category. Further, the extracted features were computed based on the automated segmentation method due to lack of precise lesion boundary for the data from Site II. Owing to the limited size of the dataset considered in this study, we did not conduct multiple statistical tests of comparisons on the features. We also did not explicitly quantify the inter-observer variability in segmentation of the dominant masses between multiple readers. One problem was the fact that we were identifying imaging markers correlated with a surrogate of outcome (*OncotypeDX*) instead of actual outcome itself- unfortunately this information was not available for the patients considered in this study.

**Concluding Remarks.** We identified a set of computer extracted image texture features on DCE-MRI that appear to be able to segregate high and low *OncotypeDX* risk scores in ER-positive breast cancers. The texture features so identified may allow for non-invasively predicting which ER-positive patients might benefit from adjuvant hormonal and chemotherapy.

## Materials and Methods

This study was approved by the institutional review board and compliant with Health Insurance Portability and Accountability Act. Written informed consent was obtained from all subjects. The experimental protocols were approved by the Case Western Reserve University Faculty of Biomedical Engineering Ethics Committee. The methods were carried out in accordance with the approved guidelines and regulations.

**Patients.** The breast DCE-MRI data were retrospectively collected from two institutions (Site I: Boston Medical Center; Site II: UH MacDonal Women's Hospital) between 2006 and 2012. All the cases were anonymised. In Site I, women patients who presented with a suspicious breast lesion on screening mammogram, then had diagnostic MRI, were recruited to a large study of MRI in the staging, diagnosis, and screening of breast cancer. In Site II, women whose pathology revealed node-negative, ER-positive invasive breast cancer who

participated in a large breast cancer case-control study were utilized for this study<sup>27</sup>. All the DCE-MRI images at 1.5T were obtained within 3–7 days after diagnostic biopsy. The lesion diagnosis for both cohorts was confirmed by ultrasound guided core needle biopsies or MRI guided biopsies, followed by the histopathologic examination of 3–10 specimens obtained by core biopsy sampling. A total of 89 patient studies were collected from the Site I, and only 17 patients with both pathology reports and available low or high *OncotypeDX* scores were included in this study. From the Site II, we acquired 101 ER-positive stage I-III female breast cancer patients. Of those, 79 patients had both associated pathology reports and available low and high *OncotypeDX* recurrence scores. Patients with intermediate *OncotypeDX* recurrence scores (18–30) were deemed to not be informative of cancer risk and hence excluded from the analysis. Three patient studies and one patient study from the Site I had two and three separate lesions respectively. All other patients only had a single lesion. For each patient, the *OncotypeDX* test was performed for the dominant mass (index lesion), hence only the index lesion in the case of the patient with multiple masses was considered. The patient selection criteria for our study are summarized in Fig. 6.

**DCE-MRI.** Of the 96 patients for whom pathology reports and *OncotypeDX* results were available, 17 breast MRIs were acquired from the Site I as multiplanar T1- and T2-weighted images on a 1.5T magnet with an 8-channel breast coil (Achieva; Philips, Best, The Netherlands). The imaging parameters for DCE-MRI were: matrix size, 252 × 286; in-plane resolution, 0.20C0.70 mm per pixel; number of temporal positions: 5–10; echo train length: 50/60; section thickness, 1.5 mmC4 mm; 4.9C7.8/2.3C4.9 [repetition time msec/echo time msec]; flip angle, 10°C30°. The remaining 79 patient studies were acquired from Site II with T1-weighted images using a 1.5T unit (MAGNETOM Avanto; Siemens, Berlin, Germany), and the imaging parameters for DCE-MRI were: matrix size, 230 × 320 C 269 × 384; in-plane resolution, 0.40C0.80 mm per pixel; number of temporal positions: 6C8; echo train length: 50; section thickness, 1.0 mmC2.5 mm; 4.7C8.1/1.5C4.5 [repetition time msec/echo time msec]; flip angle, 10°C25°. Both DCE-MRI data sets were obtained prior to, during, and after administration of 0.1 mmol/kg body weighted of gadolinium-DTPA at a flow of 4cc/second, for a total imaging duration of 5–10 minutes. Each patient study was accompanied by: i) private health information free clinical metadata containing clinical history, age range, and radiology report; ii) pathological reports containing ER-positive scale values denoting low (17–34%), moderate (34–50%), or high (50–100%); and iii) recurrence score denoting lower relapse rate and improved overall survival with adjuvant tamoxifen (<18), or the converse (>30) outcome. Table 2 summarizes the patient characteristics.

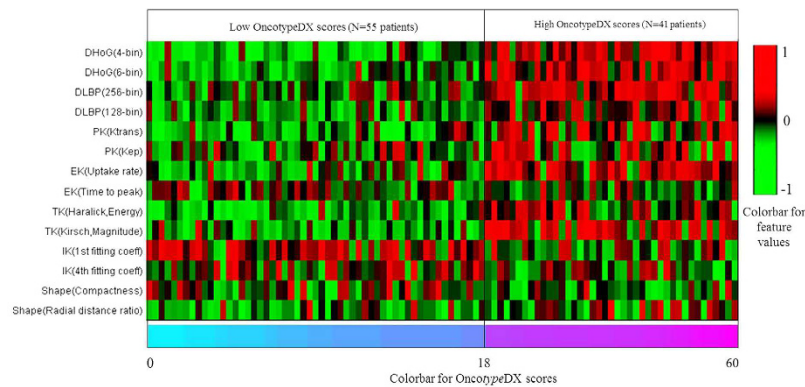
**Lesion Segmentation and Feature Extraction.** For each patient study, a representative section of the DCE-MRI volume, containing the largest diameter of the dominant mass, was chosen by a radiologist (B.N.B or D.P, both with more than 10 years of experience in the interpretation of breast MRIs) who was blinded to pathologic diagnosis. The lesion boundary was automatically delineated via an automated lesion segmentation method specifically developed and evaluated on breast DCE-MRI<sup>28</sup>. The computer derived features, including shape features, pharmacokinetics (PK), enhancement kinetics (EK), intensity kinetics (IK), TK, DHoG, and DLBP, were calculated based on the pixels enclosed by the delineated regions of interest (ROI) containing breast masses. A flowchart demonstrating the use of computerized features for lesion class discrimination is shown in Supplementary Figure S1 online. Table 3 describes the extracted features. All feature calculations were performed by using software developed in-house and was implemented using the MATLAB<sup>®</sup> programming platform (version R2013a, MathWorks, Natick, MA).

**Shape features.** Six shape features<sup>8</sup> were included: (a) area overlap ratio, (b) variance of distance ratio, (c) compactness, (d) smoothness, (e) normalized average radial distance ratio, and (f) standard deviation of normalized distance ratio. These attributes were used to measure the roundness, smoothness, spiculation, and regularity of the lesion margin.

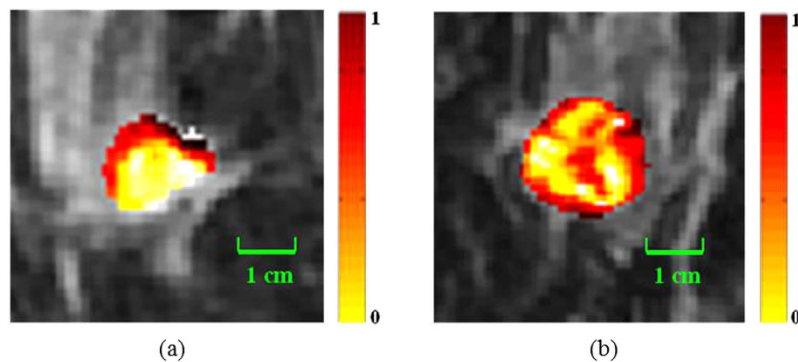
**Pharmacokinetics.** Toft's PK model<sup>29,30</sup> is most commonly used in DCE-MRI to provide a physiologic interpretation of the breast MRI images via three parameters<sup>31</sup>, i.e.,  $K^{trans}$  (the transfer constant between the plasma and tissue compartments),  $v_e$  (the extracellular extravascular volume fraction), and  $K_{ep}$  (the ratio of  $K^{trans}/v_e$ ). The PK parameters were estimated on the MRI dynamic signal enhancement curves plotted as a function of time after a bolus injection of Gd-DTPA.

**Enhancement Kinetics.** Breast lesion enhancement can be qualitatively characterized by assessing the enhancement curve obtained by plotting the signal intensity values over time after contrast injection. The mean signal intensity at each time point was calculated on the entire lesion ROI. A total of four intensity kinetic features (maximal uptake, time to peak, uptake rate, and washout rate) were computed to measure the amount and rate of contrast uptake<sup>25,26</sup>.

**Intensity Kinetics and Textural Kinetics.** A third-order polynomial was fitted to the enhancement curve to characterize its shape via a set of four model coefficients<sup>16</sup>. For each lesion, we computed five types of textural features, including Kirsch, Sobel, Haralick, and first-order textural features. Table 3 summarizes all the textural features considered in this study. The mean textural feature of lesion ROI was plotted as a function of time during the period of contrast administration. These polynomial coefficients represent the corresponding intensity and textural kinetic behavior of the lesion and represent the corresponding IK and TK features.



**Figure 1.** Heat map showing the values of best two identified features in each feature class (DHoG, DLBP, PK, EK, TK, IK, shape). The columns represent breast tumors and rows represent features. DHoG = dynamic histogram of oriented gradient; DLBP = dynamic local binary pattern; PK = pharmacokinetics; EK = enhancement kinetics; TK = textural kinetics; IK = intensity kinetics.

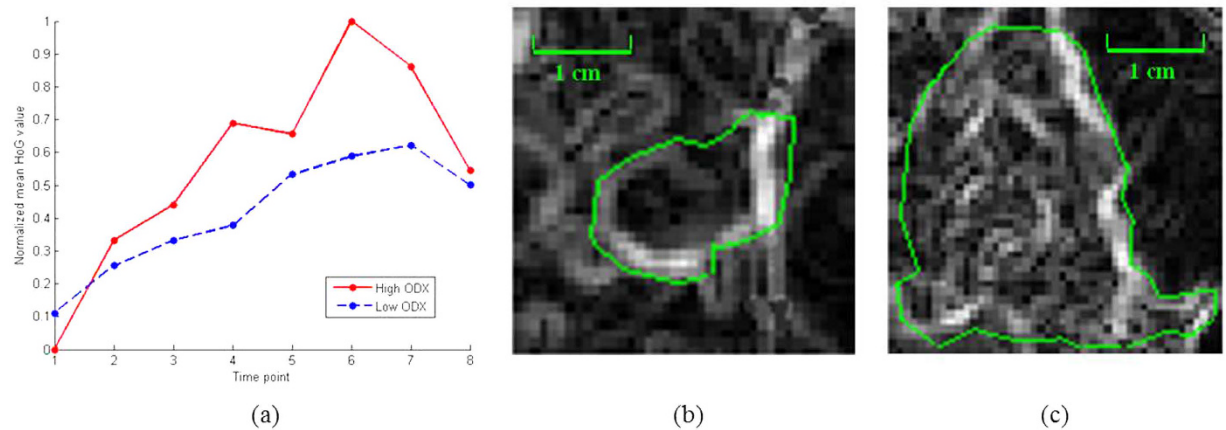


**Figure 2.** Comparison of pharmacokinetic feature ( $K^{trans}$ ) of estrogen receptor (ER)-positive breast lesions with low and high OncotypeDX recurrence scores. (a)  $K^{trans}$  in 47-year-old women with low OncotypeDX (=8), low grade ER-positive breast lesion, and (b)  $K^{trans}$  in 54-year-old women with high OncotypeDX (=58), high grade ER-positive breast lesion. The  $K^{trans}$  values are encoded in a color scale, where large values are represented in dark red and small values are represented in yellow. Note a greater heterogeneity within the high risk ER-positive breast cancers compared to low risk breast cancers.

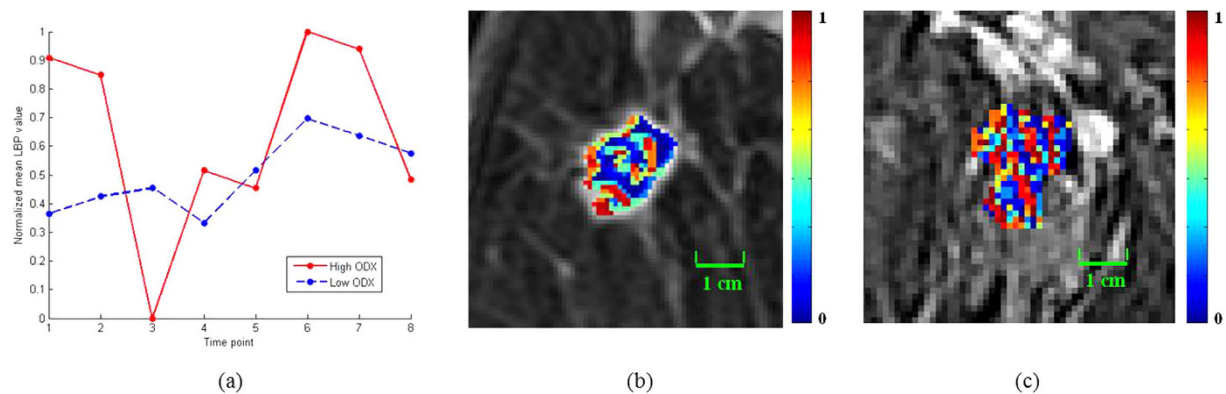
**Dynamic Histogram of Oriented Gradient Features.** We computed a multi-grid based DHoG at each phase or time point during the DCE-MRI exam. First, a gradient image at each phase was obtained via a gradient filter applied to both horizontal and vertical directions of the ROI containing the lesion. The gradient image was divided into a sequence of increasingly finer spatial grids by repeatedly doubling the number of divisions in each direction. For each grid cell, we calculated the cell histograms by counting the number of occurrences of gradient values in the histogram channels that were evenly distributed from 0 to 360 degrees. An orientation histogram was obtained by aggregating all the cell histograms. The DHoG features for the MRI time series were then obtained by averaging the orientation histograms over the course of different phases. More details on the DHoG features are described in Supplementary-A (online).

**Dynamic Local Binary Pattern Features.** Similar to the computation of the DHoG features, the lesion ROI was divided into multiple grid cells. For each pixel in the cell, we compared the pixel value to that of each of its 8 neighbors. This yielded an 8-digit binary number for the pixel under consideration. A cell histogram based on the binary numbers was then computed and normalized. An average cell histogram was calculated across phases. The DLBP features were then extracted by combining all the average cell histograms via a process of matrix concatenation. More details on the DLBP features can be obtained from Supplementary-A (online).

**Linear Discriminant Analysis based Classification via Cross-validation.** To determine computer extracted imaging features on DCE-MRI that best discriminated the low from high OncotypeDX risk categories, the LDA based classification was performed on the individual feature of each feature class (i.e., shape, PK, EK, IK, TK, DHoG, DLBP) and entire feature set containing all the feature classes (176 features in total). A LDA classifier<sup>34</sup> was trained using the extracted features to classify images with low or high OncotypeDX via an iterative 2-fold cross-validation scheme. To reduce overfitting, feature selection was performed on the entire feature set



**Figure 3.** Comparison of contrast enhancement pattern and dynamic histogram of oriented gradient (DHoG) features (4 bins) of estrogen receptor (ER)-positive breast lesions between low *OncotypeDX* recurrence score (=15), moderate grade in 49-year-old woman and high *OncotypeDX* recurrence score (=40), high grade in 64-year-old woman. (a) Normalized mean DHoG values versus time points. (b) DHoG feature map of low *OncotypeDX* at peak enhancement (7th phase, 1.5T). (c) DHoG feature map of high *OncotypeDX* at peak enhancement (6th phase, 1.5T). The green contour indicates tumor boundary. Note that the two curves have distinct enhancement patterns. Feature maps associated at peak enhancement reflect great intensity variance between two tumors.

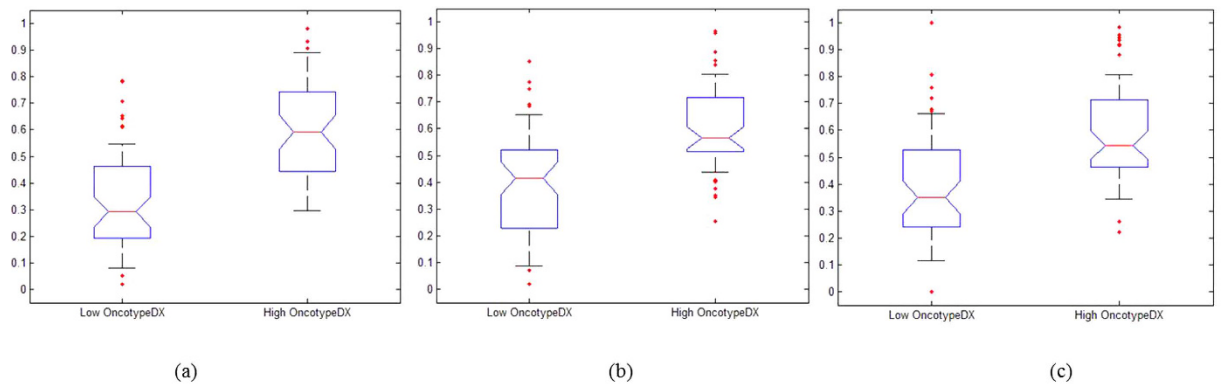


**Figure 4.** Comparison of contrast enhancement pattern and dynamic local binary pattern (DLBP) features (256 bins) of estrogen receptor (ER)-positive breast lesions between low *OncotypeDX* recurrence score (=11), low grade in 53-year-old woman and high *OncotypeDX* recurrence score (=41), high grade in 48-year-old woman. (a) Normalized mean DLBP values versus time points, and the color-coded DLBP image of (b) low *OncotypeDX* at peak enhancement (6th phase, 1.5T), and (c) high *OncotypeDX* at peak enhancement (6th phase, 1.5T). Note that the enhancement patterns vary widely in contrast uptake from time point to time point between two tumors.

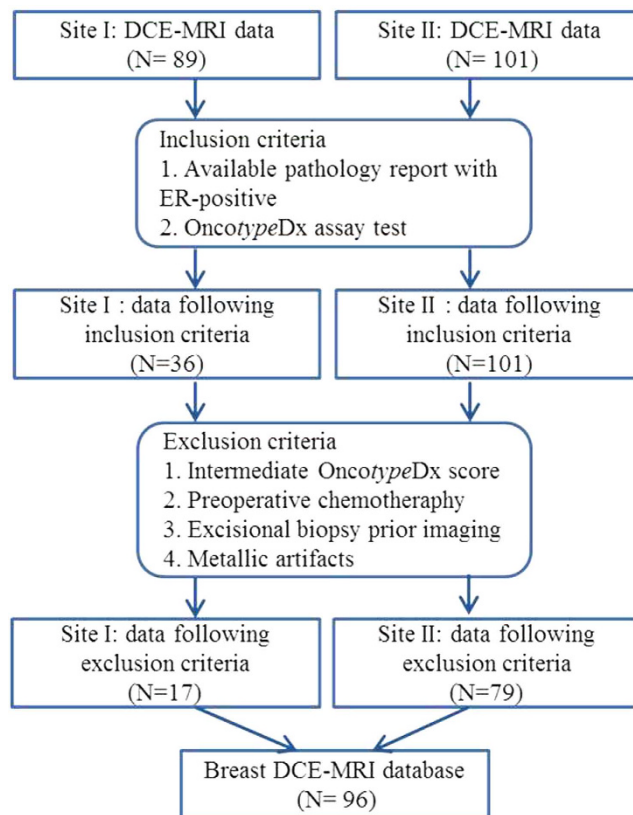
via a sequential floating forward based LDA selection method<sup>35</sup>. Further description regarding the theoretical formulation of feature selection problem and LDA classification can be obtained from Supplementary-B online. The important features were identified during the feature selection process were combined with equal weighting and used in conjunction with the LDA classifier. We assume that the condition probability density function with respect to the low and high *OncotypeDX* classes is normally distributed with equal class covariance.

**Analysis.** *Statistical Analysis.* The Student *t* test was used to verify that there was no tumor size-related bias or age-related bias between low and high *OncotypeDX* risk categories (Table 2). To confirm that our classifiers and features were robust to the choice of MRI scanners and clinical sites, we used a paired *t* test to test the null hypothesis that there were no difference in feature values between data acquired from the two sites. A post-hoc power analysis of the 95% confidence interval was performed. The Spearman's rank correlation tests measured by correlation coefficient ( $\rho$ ) were performed to determine the relationship between the computer extracted features and the low/high *OncotypeDX* risk categories. All analyses were performed by using the IBM SPSS software (version 21.0; IBM, Chicago, IL). A value of  $P < 0.05$  was considered to indicate a statistically significant difference.





**Figure 5.** Box-and-whisker plots for mean feature values of three best features corresponding to (a) dynamic histogram of oriented gradient (DHoG), (b) dynamic local binary pattern (DLBP), and (c) pharmacokinetic (PK) feature across all patient studies. The plots suggest that DHoG and DLBP have improved separability between low versus high *OncotypeDX* estrogen receptor (ER)-positive breast tumors compared to the PK features.



**Figure 6.** Flowchart of our study population with the patient inclusion and exclusion criteria.

**Stability of Classification Performance.** In the LDA classification, area under the receiver operating characteristic curve (Az), positive predictive value (PPV), negative predictive value (NPV) were used as performance measures for evaluating the discriminability of each of the individual computer extracted features. In order to assess the stability of LDA classifier, the classification was performed via a 2-fold cross validation strategy. We computed the performance measures 100 times and reported the mean values with 95% confidence interval in the results. We employed a stability measure that Parmar *et al.* used to evaluate the performance of classification methods in their recent radiomic work<sup>34</sup>. The classifier stability was empirically quantified using the relative standard deviation (RSD %), which can be defined as:

$$RSD = \frac{\sigma_{Az}}{\mu_{Az}} * 100 \quad (1)$$

where  $\mu_{Az}$  and  $\sigma_{Az}$  are the mean and standard deviation of the Az values, respectively. The higher RSD values indicate the lower stability in the classification.

**Stability and Predictive Performance.** In order to identify most accurate and highly reliable image features, we used mean values of Az and RSD as feature ranking measures. According to Parmar *et al.*'s selection criterion<sup>32</sup>, the features ranked in the top half of both measures are considered as highly accurate and reliable ones. For each feature class (i.e., shape, PK, EK, IK, TK, DHoG, DLBP), the best identified features have Az greater than the mean Az of all classifiers and RSD less than the mean RSD of all classifiers.

Further, we utilized an inverse power law model<sup>35</sup> of statistical learning to estimate the error rate associated with the classification performance on the currently available data samples. The estimation procedure comprised the following steps: (i) The dataset was divided into a training pool and a testing set via a random sampling; (ii) Ensured that the number of training samples in each set was statistically significant for calculating the power law parameters; (iii) The power law model was applied to describe the relationship between error rate and training set size:

$$err(n) = an^{-\alpha} + \varepsilon \quad (2)$$

where  $err(n)$  is the error rate for training set size  $n$ ,  $a$  is the learning rate,  $\alpha$  is the decay rate, and  $\varepsilon$  is the Bayes error. The model parameters  $[a, \alpha, \varepsilon]$  can be estimated via a constrained non-linear minimization<sup>35</sup>.

## References

- Jordan, V. C. Tamoxifen (ICI46, 474) as a targeted therapy to treat and prevent breast cancer. *Br. J. Pharmacol.* **147**, S269–S276 (2006).
- Mokbel, K. The evolving role of aromatase inhibitors in breast cancer. *Int. J. Clin. Oncol.* **7**, 279–283 (2002).
- Basavanthally, A. *et al.* Multi-field-of-view strategy for image-based outcome prediction of multi-parametric estrogen receptor-positive breast cancer histopathology: Comparison to Oncotype DX. *J. Pathol. Inform.* **2**, S1 (2011).
- Paik, S. Development and clinical utility of a 21-gene recurrence score prognostic assay in patients with early breast cancer treated with Tamoxifen. *Oncologist* **12**, 631–635 (2007).
- Paik, S. *et al.* Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J. Clin. Oncol.* **24**, 3726–3734 (2006).
- Cronin, M. *et al.* Analytical validation of the Oncotype DX genomic diagnostic test for recurrence prognosis and therapeutic response prediction in node-negative, estrogen receptor-positive breast cancer. *Clin. Chem.* **53**, 1084–1091 (2007).
- Williams, D. J. *et al.* Proliferation (Ki-67 and phosphohistone H3) and Oncotype DX recurrence score in estrogen receptor-positive breast cancer. *Appl. Immunohistochem. Mol. Morphol.* **19**, 431–436 (2011).
- Agner, S. C. *et al.* Textural kinetics: A novel dynamic contrast-enhancement (DCE)-MRI feature for breast lesion classification. *J. Digit. Imaging* **24**, 446–463 (2011).
- Ashraf, A. B. *et al.* Identification of intrinsic imaging phenotypes for breast cancer tumors: Preliminary associations with gene expression profiles. *Radiology* **272**, 374–384 (2014).
- Nie, K. *et al.* Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast MRI. *Acad. Radiol.* **15**, 1513–1525 (2008).
- Han, M. *et al.* Prognostic role of MRI enhancement features in patients with breast cancer: value of adjacent vessel sign and increased ipsilateral whole-breast vascularity. *AJR Am. J. Roentgenol.* **199**, 921–928 (2012).
- Yamamoto, S. *et al.* Radiogenomic analysis of breast cancer using MRI: a preliminary study to define the landscape. *AJR Am. J. Roentgenol.* **199**, 654–663 (2012).
- Vassiou, K. *et al.* Morphological and kinetic characteristics of dynamic contrast-enhanced MRI (DCE-MRI) correlated with histopathological factors of breast cancer: A potential prognostic role of breast MRI? *Eur. J. Radiol.* **1**, 124–132 (2009).
- Chang, Y. W. *et al.* Magnetic resonance imaging of breast cancer and correlation with prognostic factors. *Acta. Radiol.* **50**, 990–998 (2009).
- Sutton, E. J. *et al.* Breast cancer subtype intertumor heterogeneity: MRI-based features predict results of a genomic assay. *J. Magn. Reson. Imaging* **42**, 1398–1406 (2015).
- Agner, S. C. *et al.* Computerized image analysis for identifying triple-negative breast cancers and differentiating them from other molecular subtypes of breast cancer on dynamic contrast-enhanced MR images: A feasibility study. *Radiology* **272**, 91–99 (2014).
- Dalal, N. & Triggs, B. Histograms of oriented gradients for human detection. in *Computer Vision and Pattern Recognition (CVPR)*, 2005 IEEE Computer Society Conference on. 886–893 (IEEE).
- Ojala, T., Pietikäinen, M. & Harwood, D. A comparative study of texture measures with classification based on feature distributions. *Patt. Rec.* **29**, 51–59 (1996).
- Dalal, N., Triggs, B. & Schmid, C. Human detection using oriented histograms of flow and appearance. *Lecture Notes in Computer Science* **3952**, 428–441 (2006).
- Pietikäinen, M. & Zhao, G. In *Advances in Independent Component Analysis and Learning Machines* (eds Bingham, E. *et al.*) Ch. 9, 175–210 (Elsevier, 2015).
- Gardezi, S. & Faye, I. Fusion of completed local binary pattern features with curvelet features for mammogram classification. *Appl. Math. Inf. Sci.* **9**, 3037–3048 (2015).
- Ergin, S. & Kilinc, O. A new feature extraction framework based on wavelets for breast cancer diagnosis. *Comput. Biol. Med.* **51**, 171–182 (2014).
- Uematsu, T., Kasami, M. & Yuen, S. Triple-negative breast cancer: correlation between MR imaging and pathologic findings. *Radiology* **250**, 638–647 (2009).
- Harris, L. *et al.* American society of clinical oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J. Clin. Oncol.* **25**, 5287–5312 (2007).
- Chen, W. *et al.* Automatic identification and classification of characteristic kinetic curves of breast lesions on DCE-MRI. *Med. Phys.* **33**, 2878–2887 (2006).
- Bhooshan, N. *et al.* Cancerous breast lesions on dynamic contrast-enhanced MR images: Computerized characterization for image-based prognostic markers. *Radiology* **254**, 680–690 (2010).

27. Thompson, C. L. & Li, L. Association of sleep duration and breast cancer OncotypeDX recurrence score. *Breast Cancer Res. Treat.* **134**, 1291–1295 (2012).
28. Agner, S. C., Xu, J. & Madabhushi, A. Spectral embedding based active contour (SEAC) for lesion segmentation on breast dynamic contrast enhanced magnetic resonance imaging. *Med. Phys.* **40**, 032305 (2013).
29. Tofts, P. S., Berkowitz, B. & Schnall, M. D. Quantitative analysis of dynamic Gd-DTPA enhancement in breast tumors using a permeability model. *Magn. Reson. Med.* **33**, 564–568 (1995).
30. Sourbron, S. P. & Buckley, D. L. On the scope and interpretation of the Tofts models for DCE-MRI. *Magn. Reson. Med.* **66**, 735–745 (2011).
31. Newell, D. *et al.* Selection of diagnostic features on breast MRI to differentiate between malignant and benign lesions using computer-aided diagnosis: differences in lesions presenting as mass and non-mass-like enhancement. *Eur. J. Radiol.* **20**, 771–781 (2010).
32. Martinez, A. M. & Kak, A. C. PCA versus LDA. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**, 228–233 (2001).
33. Hupse, R. & Karssemeijer, N. The effect of feature selection methods on computer-aided detection of masses in mammograms. *Phys. Med. Biol.* **55**, 2893–2904 (2010).
34. Parmar, C., Grossmann, P., Bussink, J., Lambin, P. & Aerts, H. J. Machine learning methods for quantitative radiomic biomarkers. *Sci. Rep.* **5**, 13087 (2015).
35. Basavanthally, A., Viswanath, S. & Madabhushi, A. Predicting classifier performance with limited training data: Applications to computer-aided diagnosis in breast and prostate cancer. *PLoS ONE* **10**, e0117900 (2015).

## Acknowledgements

This work was made possible via grants from the National Science Foundation under award No. IIP-1248316; the National Cancer Institute of the National Institutes of Health under award Nos. R01CA136535-01, R01CA140772-01, and R21CA167811-01; the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under award No. R43EB015199-01; the National Institute for Digestive Diseases and Kidney under award No. R01DK098503; the Case Comprehensive Cancer Center Support Grant under award No. P30-CA043703; the CTSC Coulter Pilot grant from Case Western Reserve University, the V Foundation; The national natural science foundation of China under award No. 61401012; and the SRF for ROCS, SEM, China to T.W.

## Author Contributions

T.W. and A.M. were responsible for integrity of entire study. T.W. performed literature research, experimental studies and statistical analysis. B.N.B., D.P., C.L.T., H.G. and C.J. contributed to data acquisition. T.W., B.N.B., D.P., L.H. and A.M. contributed to clinical studies. All authors contributed to study concepts/study design or data analysis/interpretation, and literature research. All authors reviewed and approved the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** T.W., B.N.B., D.P., C.L.T., H.G., C.J., L.H. declare no potential conflict of interest. A.M. is a co-founder of Ibris Inc., a cancer diagnostics company. A.M. is also a PI on a NIH funded academic industrial partnership R01 grant with Siemens Corporate Research and a co-investigator on a separate NIH funded academic industrial partnership R01 grant with GE. A.M. is also stake holder in Elucid Bioimaging, a Boston based startup that is focused on novel image analysis methods in conjunction with MRI for cardiovascular applications.

**How to cite this article:** Wan, T. *et al.* A Radio-genomics Approach for Identifying High Risk Estrogen Receptor-positive Breast Cancers on DCE-MRI: Preliminary Results in Predicting OncotypeDX Risk Scores. *Sci. Rep.* **6**, 21394; doi: 10.1038/srep21394 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>