Published in final edited form as:

J Magn Reson Imaging. 2015 November ; 42(5): 1421–1430. doi:10.1002/jmri.24921.

Heterogeneity in Intratumoral Regions With Rapid Gadolinium Washout Correlates With Estrogen Receptor Status and Nodal Metastasis

Baishali Chaudhury, MS^{1,*}, Mu Zhou, MS¹, Dmitry B. Goldgof, PhD¹, Lawrence O. Hall, PhD¹, Robert A. Gatenby, MD², Robert J. Gillies, PhD², Bhavika K. Patel, MD², Robert J. Weinfurtner, MD², and Jennifer S. Drukteinis, MD²

¹Department of Computer Science and Engineering, University of South Florida, Tampa, Florida, USA

²Department of Radiology, H. Lee Moffitt Cancer and Research Institute, Tampa, Florida, USA

Abstract

Purpose—To evaluate heterogeneity within tumor subregions or “habitats” via textural kinetic analysis on breast dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) for the classification of two clinical prognostic features; 1) estrogen receptor (ER)-positive from ER-negative tumors, and 2) tumors with four or more viable lymph node metastases after neoadjuvant chemotherapy from tumors without nodal metastases.

Materials and Methods—Two separate volumetric DCE-MRI datasets were obtained at 1.5T, comprised of bilateral axial dynamic 3D T_1 -weighted fat suppressed gradient recalled echo-pulse sequences obtained before and after gadolinium-based contrast administration. Representative image slices of breast tumors from 38 and 34 patients were used for ER status and lymph node classification, respectively. Four tumor habitats were defined based on their kinetic contrast enhancement characteristics. The heterogeneity within each habitat was quantified using textural kinetic features, which were evaluated using two feature selectors and three classifiers.

Results—Textural kinetic features from the habitat with rapid delayed washout yielded classification accuracies of 84.44% (area under the curve [AUC] 0.83) for ER and 88.89% (AUC 0.88) for lymph node status. The texture feature, information measure of correlation, most often chosen in cross-validations, measures heterogeneity and provides accuracy approximately the same as with the best feature set.

Conclusion—Heterogeneity within habitats with rapid washout is highly predictive of molecular tumor characteristics and clinical behavior.

Breast tumors are heterogeneous both on genetic and histopathologic levels with intratumoral spatial variation in cellularity, angiogenesis, extravascular extracellular matrix, and areas of necrosis.¹ Generally, heterogeneity confers a poor prognosis, in part because it maximizes the probability of clones that are metastatic and/or resistant to therapy.² Cancers have been viewed as ecological systems in which molecular heterogeneity is caused by

* Address reprint requests to: B.C., 4202 E. Fowler Ave., Tampa, FL 33620. baishali@mail.usf.edu.

variations in local microenvironmental conditions largely governed by spatial and temporal changes in blood flow.³ This suggests that heterogeneity at the genetic and/or cellular levels can be correlated with tissue level heterogeneity, as seen through contrast enhancement patterns on dynamic contrast-enhanced (magnetic resonance imaging (DCE-MRI)).^{4–11} Fast progressing diseases and malignancies have been shown to be associated with highly heterogeneous enhancement patterns in DCE-MR images.¹²

The contrast enhancement pattern for a single tumor voxel is often represented through a signal intensity time curve (Fig. 1). Analysis of a representative curve for the whole tumor has gained recognition among radiologists. This analysis is often qualitative based and suffers from interobserver variability. Kinetic maps have recently been introduced to quantify the contrast enhancement pattern for each tumor voxel. Features extracted from these spatially explicit maps are used in computer-aided detection (CAD) systems to reduce the subjectivity prevalent in the current diagnosis system.

We hypothesize that the underlying cellular and molecular dynamics will be different in each tumor habitat and that clinical outcomes may be disproportionately affected by the most aggressive phenotypes within the cancer rather than the “average” intratumoral phenotype. Our goal was to identify the most predictive tumor habitats and correlate the heterogeneity within each habitat to key clinical and prognostic features.

MATERIALS AND METHODS

Dataset Acquisition

An Institutional Review Board (IRB) and Health Insurance Portability and Accountability Act (HIPAA)-compliant retrospective review was performed on all Breast Imaging and Reporting Data System (BI-RADS) 5 and 6 DCE-MRI reports from a single institution from January 1, 2010 to July 1, 2014. A database was constructed by obtaining data of consecutive clinical stage II and III breast cancer patients, with tumors ≥ 2.0 cm, who did not undergo any treatment for their breast cancer prior to their initial DCE-MRI.

Consecutive patients from the database that satisfied the necessary criteria were selected for the two tasks of estrogen receptor (ER) status classification, and viable lymph node status classification after neoadjuvant chemotherapy. No additional information was known about the patients apart from their ER status and lymph node status at final surgery after neoadjuvant chemotherapy when selecting the two groups for task classification. Images from seven patients were utilized for both datasets, as the images for analysis were applicable for both tasks. For classification of ER status, the dataset included images of 38 patients (20 ER-positive and 18 ER-negative) with a histopathologic diagnosis of invasive ductal or invasive lobular breast carcinoma. For the task of ER classification, 18 consecutive ER-negative cases were obtained; attention was then turned to the first 20 consecutive ER-positive cases. ER-negative cases that fulfilled our criteria were the limitation. For ER status classification, the cohort consisted of women ages 31–74 years of age, with a mean age of 52.7 years and median of 52 years. Thirty-five patients had a diagnosis of invasive ductal carcinoma and three patients had a diagnosis of invasive lobular carcinoma. ER status classification was determined at core biopsy with a cutoff of 10% for ER-positive tumors.

For viable lymph node classification, our pilot data consisted of 34 patients who underwent neoadjuvant chemotherapy prior to surgery. For the task of lymph node classification, 16 consecutive cases of patients with no viable lymph node metastases were obtained; attention was then turned to the first 18 consecutive cases with four or more positive lymph node metastases. There were 16 patients with no viable lymph node metastases (labeled as class NoLN) and 18 patients with four or more viable lymph node metastases (labeled as class FourOrMoreLN). The presence of viable lymph node metastases was determined at final pathology, after neoadjuvant chemotherapy. We reasoned that the presence of viable lymph node metastases, after neoadjuvant chemotherapy, is a marker for increased risk of subsequent development of metastases, and thus, a poorer prognosis. For lymph node status classification, the cohort consisted of women ages 39–74 years of age, with a mean age of 53.7 years and median age of 52 years. Thirty-three patients had a diagnosis of invasive ductal carcinoma and one had a diagnosis of invasive lobular carcinoma.

Diagnostic breast MRI was performed using a GE 1.5T Optima 450w MRI and Sentinelle double-breast coil. The breast DCE-MRI technique included bilateral axial dynamic 3D T_1 -weighted fat-suppressed gradient recalled echo-pulse sequences obtained before and after gadolinium-based contrast administration. Dynamic images were obtained every 90–110 seconds adjusted for breast thickness for a total of six acquisition times. Slice thickness ranged from 1.1–2.2 mm. Gadolinium based contrast agents were used at a concentration of 0.1 mmol/kg. From January 1, 2010 to July 1, 2011, gadopentetate dimeglumine (Magnevist, Bayer HealthCare Pharmaceuticals, Berlin, Germany) was used. From July 2, 2011 to July 1, 2014, gadobutrol (Gadavist, Bayer HealthCare Pharmaceuticals) was used. For the purposes of this study, analysis was performed exclusively on the T_1 fat-saturated pre- and postcontrast sequences. The pixel resolution of the images varied from 1.3–1.7 pixels/mm, and 1.5 pixels/mm was found to be the most common pixel resolution for all the images. The bit depth was 16 bits/pixel. The T_1 acquisition parameters were otherwise the same over time. In order to prevent inaccuracies when computing textures due to resolution inconsistency,¹³ all images were reinterpolated¹⁴ to standardize the pixel resolution to 1.5 pixels/mm. For each patient a single breast-imaging radiologist (J.D., 9 years of experience) selected a representative image slice with the representative tumor cross-section, from the DCE-MRI volumetric data that was then utilized for analysis. For only the lymph node dataset, we had a second radiologist (R.W., 6 years of experience) select the representative tumor cross-section, without prior knowledge of the previously chosen slices. Out of 34 cases, 13 cases had different representative slices chosen by the second radiologist. Among these 13 cases, eight cases had almost the same tumor size (visually) and for the other five cases the slices chosen by the second radiologist were visually larger than the previously chosen slices.

For patients with multicentric or multifocal malignancy, the largest tumor was chosen for analysis. We analyzed four contrast based habitats: regions with rapid-initial-uptake, slow-initial-uptake, persistent-delayed-washout, and rapid-delayed-washout for the classification of ER-positive from ER-negative tumors and tumors with four or more lymph nodes from tumors without lymph node metastases.

Tumor Segmentation and Extraction of Contrast-Based Habitats

The precontrast image was subtracted either from the first or second postcontrast image (where the peak intensity was reached) on which a two class Otsu segmentation¹⁵ was performed to identify the tumor pixels. Dilation¹⁶ (using a square of size 5 as the structuring element) followed by connected components analysis gave the final tumor mask. Figure 2a–h demonstrates the segmentation process.

We consider the tumor as a coalition of different “habitats” and seek to identify the imaging characteristics and heterogeneity of each and its relative contribution to the clinical behavior of the entire breast cancer. Through quantitative image analysis, we examined the presence of differing tumor habitats defined by variations in initial and delayed contrast patterns within the tumor. We extracted four contrast-based tumor habitats by grouping similar tumor pixels that showed similar contrast enhancement patterns on DCE-MR images, both in the initial phase and the delayed phase. We used the BI-RADS intensity curve just as a platform for analysis, but did not base the habitat segmentation on it. Hence, instead of applying fixed thresholding for segmentation we used the two-class Otsu segmentation, which is a clustering-based thresholding. Otsu segmentation assumes that the image contains two classes of pixels having a bimodal intensity histogram and computes the optimal threshold to divide the image into two classes. To group pixels based on their initial uptake contrast patterns, we conducted a pixelwise subtraction for the tumor pixels, between their intensity in precontrast and their peak intensity reached either in the first or second postcontrast image. A two-class Otsu segmentation on the resultant subtracted image divided the tumor into two subregions. The subregion with a higher average intensity was labeled as rapid-initial-uptake and the other was labeled as slow-initial-uptake. For grouping pixels based on their washout patterns we conducted a pixelwise subtraction only for the tumor pixels between pixel's peak intensity (either the first or second postcontrast image) and its intensity in the last postcontrast image. The subtracted tumor region was divided into two subregions using two class Otsu segmentation. The subregion with a higher average intensity was labeled as persistent-delayed-washout and the other was labeled as rapid-delayed-washout. Figure 2i–l demonstrate the above process.

Multiparametric Texture Feature Extraction

To quantify initial and delayed contrast enhancement patterns for each tumor pixel (Fig. 1) we computed initial enhancement (IE) and postinitial enhancement (PIE) values for each tumor pixel as described previously.^{14,17} In order to generate IE and PIE kinetic maps, the IE and PIE values for all the tumor pixels were scaled between 1 and G-1 gray levels globally (across all the patients). For global scaling, the lowest PIE value (LowPIE) and highest PIE value (HighPIE)¹⁷ corresponded to the 0.5th and 0.95th percentile of all the PIE values, respectively. The same approach was used to calculate the lowest IE (LowIE) and highest IE (HighIE)¹⁷ values. The values for LowIE, HighIE, LowPIE, and HighPIE were recorded and are presented in the Results section. We then extracted texture features from these maps and used them to measure image heterogeneity by assessing the relationships among the image gray levels using a variety of texture features.

We extracted 22 gray level co-occurrence matrix (GLCM),^{18,19} 11 gray level run length matrix (GLRL),²⁰ and 76 local binary pattern histogram Fourier (LBP-HF)²¹ features from both IE and PIE kinetic maps of the tumors. GLCM features calculate the spatial relationship between different image gray levels, while GLRL features compute the occurrence and the distribution of the same gray level. LBP features measure local image heterogeneity by comparing each image pixel to neighboring pixels. For each tumor pixel, texture features were computed by considering a window of size $m \times m$ around it. This window size is also known as the scale for texture computation. We obtained a final feature set of 109 features for each habitat, by averaging the pixelwise features over the habitat. The amount of information captured by texture features depends upon the scale of texture computation ($m \times m$), the quantization levels (Q) used for computing GLRL and GLCM, and the number of gray levels (G) used for scaling IE and PIE values. Mixing features from “different levels” has the potential to provide useful information to a classifier. To capture the maximum information out of the texture features from each habitat, it is important to find the optimal values for these parameters for each habitat. In order to have a comprehensive comparison of the four contrast-based habitats, these three parameters were varied to generate different sets of parameters (PS). Corresponding to these PS(s), feature sets were extracted. Table 1 summarizes the ranges for these parameters. Figure 3 depicts the multiparametric texture feature extraction process.

Feature Selection and Classification

The features were normalized into $[-1,1]$ before feature selection and classification. The number of features is larger than the number of instances, so the use of feature selection prior to classification is mandatory in order to avoid overfitting and creating spurious classification models. We employed two feature selection methods, the correlation-based feature subset selection (CFS) algorithm²² and a wrapper-based feature subset selection.²³ CFS is a filter-based method that builds the feature subset such that the features are uncorrelated with each other, but “positively” correlated to the class. Wrappers uses cross-validation to evaluate the merit of a feature subset using a particular classifier. Three classifiers were evaluated: naive Bayes,²⁴ decision trees²⁵ and support vector machines (SVM), via the LIBSVM package, in the Weka tool.^{26,27} The naive Bayes classifier is a simple probabilistic classifier that calculates the probability of a test instance belonging to a particular class using prior probabilities and likelihoods. SVM is a supervised learning method, which typically transforms the data into linearly separable data by projecting it onto a high-dimensional feature space to find a maximum margin hyperplane. Decision trees recursively use the attribute with maximum normalized information gain from the training data to classify the data samples. We used six feature selector-classifier pairs or “meta-classifiers” (Table 2). We conducted feature selection and classification fold-wise, in a leave-one-out (L-O-O) fashion.⁴¹ We divided n samples into n folds and on each fold we did feature selection on the $n-1$ training samples, built the classifier using the selected features, and then tested on the 1 left out test sample. Since we worked with a large number of texture features (109 features) and conducted feature selection in every fold of the data, it is instructive to form a feature subset that appears most frequently across folds of the different feature selector-classifier pairs. To check this repeatability, we applied selection criteria exclusively to the best results. We formed the subset of only those features that were selected

in more than 65% of the folds (ie, 25 folds of the 38 folds for ER prediction and 22 folds of the 34 folds for lymph node prediction) during the fold-wise feature selection, and we refer them as the most common features.

The feature sets were evaluated in terms of average class accuracy (Avg Acc), sensitivity (Sens), and specificity (Spec),²⁸ all expressed as percent. In order to construct receiver operating characteristic (ROC) curves, an experiment using bootstrapping 10 times with different random seeds was done to get multiple points on an ROC curve. Forward feature selection with wrappers was done for generating the ROC curves. We also calculated the kappa statistic against a random classifier for the best classifier and averaged value over the 10 trials.

RESULTS

The values for [LowPIE, HighPIE]⁴⁰ were found to be $[-0.73, 1.59]$ and the values for [LowIE, HighIE] were found to be $[0.12, 34.68]$, for the dataset used for ER status prediction. For the dataset for lymph node prediction the values for [LowPIE, HighPIE] were found to be $[-1, 1.45]$ and the values for [LowIE, HighIE] were found to be $[0.011, 9.04]$. In Table 3 we summarize the best results from each meta-classifier for ER status classification (majority class is 52.6%). The ROC curves for three classifiers for ER status classification are shown in Fig. 4. An average AUC of 0.836 and kappa statistic of 0.7 was obtained for naive Bayes on the ER status dataset.

In Tables 4 and 5 we summarize the best results from each meta-classifier for lymph nodes classification (FourOr-MoreLN vs. NoLN) using the slices chosen by the first (J.D.) and the second (R.W.) radiologist, respectively (majority class is 52.9%). Comparing Tables 4 and 5, it was found that the results on the slices chosen by the second radiologist were a bit lower than those obtained on slices chosen by the first radiologist. The highest accuracy of 88.89% was achieved using the slices chosen by the first radiologist while when using slices chosen by the second radiologist the highest accuracy was 82.9% (two extra errors were made). In Fig. 5 we compare the ROC curves for both the sets of slices for lymph node classification. For slices chosen by the first radiologist an average AUC of 0.886 for J48 was obtained and for slices chosen by second radiologist the average AUC was 0.8. The kappa statistic indicated that the results from our classifiers were well above those from random guessing of the classes, with kappa values as high as 0.77.

To analyze the effect of differing slice selection, we trained on the set of slices chosen by the second radiologist (R.W.) and tested on those chosen by the first radiologist (highest average class accuracy of 80.5% using wrappers feature selection and J48), and vice versa (highest average class accuracy of 72.5% using wrappers feature selection and SVM). For the first experiment, the number of extra errors (compared to the testing on the training dataset itself) made on the cases with different slices was only 2, but for the second experiment the number of extra errors made on the cases with different slices was 9. This indicates that training on the second set of slices was better and did not overfit as much as the first set of slices did. One of the possible reasons can be that there were five slices that were bigger in the second set than in the first case, resulting in slightly more accurate features.

In order to understand the efficacy of the features we conducted the above experiment using only LBP-HF features. The results using LBP-HF features were more stable across the two train and test sets than the results which also used GLCM and GLRL features. This can be attributed to the fact that GLCM and GLRL depend on the size of the image. However, for LOO-CV the results were worse when only LBP-HF features were used.

The most common features (based on the criteria explained in the Materials and Methods), for ER status and lymph node classification are summarized in Table 6. We found that the GLCM texture feature “Information measure of correlation”²⁹ was the most repeatedly selected feature for both ER status and lymph node status classification. This GLCM feature represents the correlation between gray levels identified via texture analysis within the tumor habitat and is inversely proportional to heterogeneity. The class means of ER-positive and ER-negative tumors, for the information measure of correlation feature, were -0.01 and -0.64 , with standard deviations of 0.47 and 0.24 , respectively. The class means of NoLN category and FourOrMoreLN category, for the same feature, were -0.36 and -0.66 , with standard deviations of 0.38 and 0.22 , respectively. Using this single feature we achieved an average class accuracy of 88.89% (using slices chosen by J.D.) for lymph nodes and 84.4% for ER status using LOO-CV using the J48 classifier. However, a statistical difference between the class means was not found between the two classes (for both the classification tasks) for this feature.

DISCUSSION

Our analysis is a novel approach to the study of heterogeneity in breast cancer. Rather than whole tumor analysis, we divided each breast tumor into differing habitats based on the blood flow dynamics following gadolinium injection. We quantified the heterogeneity within each habitat to correlate that with two clinically important global tumor features of ER status and relatively treatment resistant viable metastases to axillary lymph nodes.

Previous work has identified that contrast enhancement patterns correlate with clinical and histologic features, such as axillary lymph node metastases^{30–34} and ER status.^{35–41} However, prior studies based on kinetic maps have extracted features from the whole tumor.^{17,39,42,43} This approach was based on the tacit assumption that tumors are heterogeneous, but well-mixed. Thus, features averaged across the entire tumor and regions with different imaging and genotypic characteristics contribute to the same extent in the final feature computation. Previous attempts^{11,12,15,44–51} to identify tumor subregions have mainly concentrated on benign vs. malignant classification and did not provide a comprehensive comparison of different habitats. Hence, these methods have shown little success in identifying the most predictive habitats.

Our results have three potential implications. First, heterogeneity quantified through textural analysis of MRI has the potential to serve as a surrogate for heterogeneity at the cellular and genetic level. Second, the degree of intratumor heterogeneity can vary within habitats and within the tumor itself. Finally, heterogeneity within tumor habitats with rapid delayed washout may contribute disproportionately to the clinical phenotype of the tumor. The GLCM texture feature, “information measure of correlation” (extracted from the rapid

delayed washout habitat), which measures heterogeneity, was found to be the single best texture feature from all the data that enabled nearly as good predictive accuracy as feature selection during cross-validation. Thus, this feature is strongly discriminative. However, a bigger dataset and further analysis is required for validation.

The role of regions with rapid vascular washout in both ER expression and nodal metastases within breast cancer suggests further investigation is warranted. These findings suggest that either the rapid washout vasculature is promoted by a particularly important cancer cell phenotype (niche engineering) or the pattern of blood flow selects for a particularly aggressive cancer cell phenotype.

Our study does not allow us to identify the cause(s) of heterogeneity within each subregion. Generally, the findings could be due to the presence of different cell populations within a subregion including mixtures of tumor and host parenchymal cells (ie, fibrosis), as infiltration by stromal cells has been linked with delayed washout and rapid wash-out.⁵² Alternatively, blood flow could vary temporally or over small spatial scales. Interestingly, the latter environmental conditions would tend to select for generalist cancer phenotypes, that is, cells that can adapt to a wide range of environments. Such populations would probably be more likely to survive in distant metastatic sites.

Our study represents a preliminary test of our underlying hypothesis and has several limitations including a retrospective design. Furthermore, a single slice was selected to be representative of the tumor volume. The experiments on differing slices for lymph nodes classification demonstrate the effect of representative slice selection, although we also attribute this to the effect of small sample size. Thus, use of the volumetric tumor for analysis is warranted. This will ensure adequate representation of the whole tumor volume, thus removing the ambiguity from choosing a representative slice. Hence, volumetric analysis would present a more powerful analysis. The results on differing slices also show some of the weaknesses of the texture features utilized here. Lymph node status was also analyzed at definitive surgery, after all patients completed neoadjuvant treatment. Although we realize that this may be imperfect, it can be inferred that patients who do not respond well to designated therapy and that have greater than four nodes with viable tumor at final surgery have aggressive tumor phenotypes and are at greater risk for distant metastases. Tumor response to neoadjuvant chemotherapy in patients with breast cancer has been shown to correlate with disease-free and overall survival rates.^{53,54} Adjusting for the confounding effect of neoadjuvant regimens as well as additional patient prognostic factors was beyond the scope of this analysis. Direct histologic correlation within tumor habitats is not possible, as enhancement kinetics can only be analyzed in vivo. Furthermore, the analyzed cohort is relatively small and our results will need to be confirmed in larger retrospective and prospective studies.

In conclusion, we demonstrated that clinical imaging can be used to quantify intratumoral heterogeneity and that habitats of breast tumors displaying rapid washout enhancement may contribute disproportionately to the phenotypic clinical behavior of the tumor. This work has the potential to impact patient outcomes following a diagnosis of breast cancer by automating quantifiable characteristics of tumor heterogeneity that contribute to patient

prognosis and response to treatment. This information could prove to be invaluable in determining patient treatment regimens and stratifying patient prognosis for surveillance regimens. It also has the potential to dramatically improve the quality of life for patients who are diagnosed with indolent cancer, those with tumors less susceptible to distant spreading, potentially allowing for less toxic treatments and/or surgery.

Acknowledgments

Contract grant sponsor: Department of Diagnostic Imaging, Moffitt Cancer Center “Breast Cancer Heterogeneity: Quantitative Analysis of MRI Features of Breast Cancer.”

References

1. Davnall F, Yip CS, Ljungqvist G, et al. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights Imaging*. 2012; 3:573–589. [PubMed: 23093486]
2. Kuukasjärvi T, Karhu R, Tanner M, et al. Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res*. 1997; 57:1597–1604. [PubMed: 9108466]
3. Zhou M, Hall L, Goldgof D, et al. Radiologically defined ecological dynamics and clinical outcomes in glioblastoma multiforme: preliminary results. *Transl Oncol*. 2014; 7:5–13. [PubMed: 24772202]
4. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. 2012; 366:883–892. [PubMed: 22397650]
5. Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res*. 2012; 72:4875–4882. [PubMed: 23002210]
6. Ng CK, Pemberton HN, Reis-Filho JS. Breast cancer intratumor genetic heterogeneity: causes and implications. *Expert Rev Anticancer Ther*. 2012; 12:1021–1032. [PubMed: 23030222]
7. Soussan M, Orlhac F, Boubaya M, et al. Relationship between tumor heterogeneity measured on fdg-pet/ct and pathological prognostic factors in invasive breast cancer. *PloS one*. 2014; 9:e94017. [PubMed: 24722644]
8. Shah SP, Morin RD, Khattri J, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature*. 2009; 461:809–813. [PubMed: 19812674]
9. Leth-Larsen R, Terp MG, Christensen AG, et al. Functional heterogeneity within the cd44 high human breast cancer stem cell-like compartment reveals a gene signature predictive of distant metastasis. *Mol Med*. 2012; 18:1109–1121. [PubMed: 22692575]
10. Glaser S, Niemann U, Preim B, Spiliopoulou M. Can we distinguish between benign and malignant breast tumors in DCE-MRI by studying a tumor's most suspect region only? *IEEE 26th International Symposium on Computer-Based Medical Systems (CBMS)*. 2013:77–82.
11. Mahrooghy, M.; Ashraf, AB.; Daye, D., et al. *Medical Image Computing and Computer-Assisted Intervention-MICCAI*. Springer; Berlin: 2013. Heterogeneity wavelet kinetics from DCE-MRI for classifying gene expression based breast cancer recurrence risk.; p. 295-302.
12. Yang X, Knopp MV. Quantifying tumor vascular heterogeneity with dynamic contrast-enhanced magnetic resonance imaging: a review. *J Biomed Biotechnol*. 2011:732–848.
13. Chen W, Giger ML, Li H, Bick U, Newstead GM. Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. *Magn Reson Med*. 2007; 58:562–571. [PubMed: 17763361]
14. Chaudhury B, Zhou M, Goldgof DB, et al. Using features from tumor subregions of breast DCE-MRI for estrogen receptor status prediction. *2014 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. 2014:2624–2629.
15. Otsu N. A threshold selection method from gray-level histograms. *Automatica*. 1975; 11:285–296. 23–27.
16. Sonka M, Hlavac V, Boyle R. *Image processing, analysis, and machine vision*. Cengage Learn. 2014 [Epub ahead of print].

17. Karahaliou A, Vassiou K, Arikidis N, Skiadopoulos S, Kanavou T, Costaridou L. Assessing heterogeneity of lesion enhancement kinetics in dynamic contrast-enhanced MRI for breast cancer diagnosis. *Br J Radiol*. 2014; 83:296–309. [PubMed: 20335440]
18. Uppuluri, A. GLCM texture features. Nov 20. 2008 Available at <http://www.mathworks.com/matlabcentral/fileexchange/22187-glcm-texture-features>
19. Haralick R, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*. 1973; 3:610–621.
20. Wei, X. Gray level run length matrix toolbox v1. 0, Software. Beijing Aeronautical Technology Research Center; Nov 12. 2007 Available at <http://www.mathworks.com/matlabcentral/fileexchange/17482-gray-level-run-length-matrix-toolbox>
21. Ahonen T, Matas J, He C, Pietikainen M. Rotation invariant image description with local binary pattern histogram Fourier features. In: *Image analysis*. Berlin: Springer. 2009:61–70.
22. Hall, MA. PhD thesis. University of Waikato; 1999. Correlation-based feature selection for machine learning..
23. Kohavi R, John GH. Wrappers for feature subset selection. *Artif Intel*. 1997; 97:273–324.
24. John, GH.; Langley, P. *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann; Burlington, MA: 1995. Estimating continuous distributions in Bayesian classifiers.; p. 338–345.
25. Quinlan, JR. C4. 5: programs for machine learning. Vol. 1. Morgan Kaufmann; Burlington, MA: 1993.
26. Chang CC, Lin CJ. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*. 2011; 2:27.
27. El-Manzalawy, Y.; Honavar, V. Wlsvm: Integrating libsvm into weka environment. 2005. Available at <http://www.cs.iastate.edu/yasser/wlsvm>
28. Powers DM. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *J Machine Learn Technol*. 2011; 2:37–63.
29. Linfoot E. An informational measure of correlation. *Inform Control*. 1957; 1:85–89.
30. Li C, Meng S, Yang X, Wang J, Hu J. The value of t2* in differentiating metastatic from benign axillary lymph nodes in patients with breast cancer — a preliminary in vivo study. *PloS One*. 2014; 9:e84038. [PubMed: 24454715]
31. Bhooshan N, Giger ML, Jansen SA, Li H, Lan L, Newstead GM. Cancerous breast lesions on dynamic contrast-enhanced MR images: computerized characterization for image-based prognostic markers. *Radiology*. 2010; 254:680–690. [PubMed: 20123903]
32. Tuncbilek N, Tokatli F, Altaner S, et al. Omurlu K, Temizoz O. Prognostic value DCE-MRI parameters in predicting factor disease free survival and overall survival for breast cancer patients. *Eur J Radiol*. 2012; 81:863–867. [PubMed: 21398061]
33. Jensen LR, Garzon B, Heldahl MG, et al. Diffusion-weighted and dynamic contrast-enhanced MRI in evaluation of early treatment effects during neoadjuvant chemotherapy in breast cancer patients. *J Magn Reson Imaging*. 2011; 34:1099–1109. [PubMed: 22002757]
34. Rahbar H, Partridge SC, Javid SH, Lehman CD. Imaging axillary lymph nodes in patients with newly diagnosed breast cancer. *Curr Probl Diagnost Radiol*. 2012; 41:149–158.
35. Jinguji M, Kajiya Y, Kamimura K. Rim enhancement of breast cancers on contrast enhanced MR imaging: relationship with prognostic factors. *Breast Cancer*. 2006; 13:64–73. e al. [PubMed: 16518064]
36. Matsubayashi R, Matsuo Y, Edakuni G, Satoh T, Tokunaga O, Kudo S. Breast masses with peripheral rim enhancement on dynamic contrast-enhanced MR images: correlation of MR findings with histologic features and expression of growth factors 1. *Radiology*. 2000; 217:841–848. [PubMed: 11110952]
37. Teifke A, Behr O, Schmidt M, et al. Dynamic MR imaging of breast lesions: Correlation with microvessel distribution pattern and histo-logic characteristics of prognosis 1. *Radiology*. 2006; 239:351–360. [PubMed: 16569783]
38. Chen JH, Baek H, Mehta R, Carpenter P, Nalcioğlu O, Su M. MR imaging features of triple negative breast cancer. *Breast J*. 2013; 19:643–649. [PubMed: 24015869]

39. Agner S, Rosen M, Englander S, et al. Distinguishing molecular sub-types of breast cancer based on computer-aided diagnosis of DCEMRI. International Society for Magnetic Resonance in Medicine Annual Meeting. 2010; 2490
40. Golden DI, Lipson JA, Telli ML, Ford JM, Rubin DL. Dynamic contrast-enhanced MRI-based biomarkers of therapeutic response in triple-negative breast cancer. *J Am Med Inform Assoc*. 2013; 20:1059–1066. [PubMed: 23785100]
41. Chen JH, Baek HM, Nalcioğlu O, Su MY. Estrogen receptor and breast MR imaging features: a correlation study. *J Magn Reson Imaging*. 2008; 27:825–833. [PubMed: 18383260]
42. Milenkovic J, Hertl K, Kosir A, Zibert J, Tasic JF. Characterization of spatiotemporal changes for the classification of dynamic contrast-enhanced magnetic-resonance breast lesions. *Artif Intell Med*. 2013; 58:101–114. [PubMed: 23548472]
43. Chaudhury B, Hall LO, Goldgof DB, Gatenby RA, Gillies R, Drukteinis JS. New method for predicting estrogen receptor status utilizing breast MRI texture kinetic analysis. *SPIE Medical Imaging, International Society for Optics and Photonics*. 2014:90351V–90351V.
44. Chaudhury B, Zhou M, Hall LO, et al. Using features from tumor sub-regions of breast DCE-MRI for estrogen receptor status prediction. *Press Systems, Man and Cybernetics*. 2014
45. Pickles MD, Lowry M, Manton DJ, Gibbs P, Turnbull LW. Role of dynamic contrast enhanced MRI in monitoring early response of locally advanced breast cancer to neoadjuvant chemotherapy. *Breast Cancer Res Treat*. 2005; 91:1–10. [PubMed: 15868426]
46. Twellman T, Lichte O, Nattkemper TW. An adaptive tissue characterization network for model-free visualization of dynamic contrast-enhanced magnetic resonance image data. *IEEE Transactions on Medical Imaging*. 2004; 24:1256–1266.
47. Mohajer, M. PhD thesis. 2012. Cluster analysis of the signal curves in perfusion DCE-MRI datasets.. lmu
48. Eyal E, Badikhi D, Furman-Haran E, Kelcz F, Kirshenbaum KJ, Degani H. Principal component analysis of breast DCE-MRI adjusted with a model-based method. *J Magn Reson Imaging*. 2009; 30:989–998. [PubMed: 19856419]
49. Chen L, Choyke PL, Chan TH, Chi CY, Wang G, Wang Y. Tissue-specific compartmental analysis for dynamic contrast-enhanced MR imaging of complex tumors. *IEEE Transactions on Medical Imaging*. 2011; 30:2044–2058. [PubMed: 21708498]
50. Eyal E, Degani H. Model-based and model-free parametric analysis of breast dynamic-contrast-enhanced MRI. *NMR Biomed*. 2009; 22:40–53. [PubMed: 18022997]
51. Agner SC, Xu J, Fatakdawala H, et al. Segmentation and classification of triple negative breast cancers using DCE-MRI. *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium*. 2009:1227–1230.
52. Cheng L, Li X. Breast magnetic resonance imaging: kinetic curve assessment. *Gland Surg*. 2013; 2:50–53. [PubMed: 25083457]
53. Rastogi P, Anderson SJ, Bear HD, et al. Preoperative chemotherapy: updates of National Surgical Adjuvant. *J Clin Oncol*. 2008; 26:778–785. [PubMed: 18258986]
54. Raina V, Kunjahari M, Shukla NK, et al. Outcome of combined modality treatment including neoadjuvant chemotherapy of 128 cases of locally advanced breast cancer: data from a tertiary cancer center in northern India. *Indian J Cancer*. 2011; 48:80. [PubMed: 21248440]

FIGURE 1.

Signal intensity time/kinetic curve for a particular voxel. This curve shows the contrast enhancement pattern of a tumor voxel in T_1 MRI, fat-suppressed images following injection of gadolinium. Initial enhancement (IE) and postinitial enhancement (PIE) kinetic maps are generated by quantifying the initial and the delayed phase for each pixel within the tumor, respectively.

FIGURE 2.

Demonstration of tumor segmentation and extraction of contrast based habitats. (a) Precontrast image; (b) first post-contrast image; (c) second postcontrast image; (d) last postcontrast image; (e) precontrast image subtracted from first or second postcontrast image (where the peak intensity is achieved) to highlight the tumor; (f) 2 class Otsu segmentation performed on the entire image in e; (g) dilation performed on image f; (h) connected component analysis applied on image g to retain the largest blob which represents the tumor mask M ; (i) retaining only those pixels in image e which are within tumor mask in h; (j) 2 class Otsu segmentation within the tumor in image i to obtain habitats based on initial uptake contrast enhancement pattern slow-initial-uptake (shown in the image using lower intensity) and rapid-initial-uptake (shown in the image using higher intensity and represents tumor habitat which undergo high or rapid initial contrast uptake); (k) tumor in first or second postcontrast image subtracted from tumor in last postcontrast image; (l) 2 class Otsu segmentation within the tumor in image k to segment tumor based on delayed postcontrast pattern rapid-delayed-washout (shown in the image using lower intensity and represents tumor habitat with rapid washout) and persistent-delayed-washout (shown in the image using higher intensity).

FIGURE 3.

Schematic for multiparametric texture feature extraction from PIE (postinitial enhancement) kinetic maps of contrast-based habitats (same was done for initial enhancement kinetic maps). Different feature sets (FS) corresponding to different parameter sets (PS), were extracted from different habitats. The multiparametric feature sets were then evaluated independently using six meta-classifiers, C1 to C6.

FIGURE 4.
ROC curve for ER-negative class.

FIGURE 5.

A comparison of ROC curves for tumors with four or more lymph nodes (FourOrMoreLN) for the slices chosen by the first radiologist and the slices chosen by the second radiologist.

TABLE 1

Feature Extraction Parameters and Their Ranges

Parameter	Description	Range
G	Number of gray levels for scaling IE and PIE values	64,128,256 (Total=3)
Q	Number of quantization levels for constructing GLCM and GLRL	8,16,32,64 (Total =4)
m	Window size ($m \times m$) for texture computation	7,9,11,13,15,17 (Total= 6)
PS per habitat = $3 \times 4 \times 6 = 72$		

TABLE 2

Different Meta-Classifiers and Their Descriptions

Meta-classifier	Feature selector	Classifier	Description
C1	Wrappers	Naive Bayes	Wrappers: Best first Forward selection, LOO-CV to find features
C2	Wrappers	J48	Confidence Factor= 0.5 for J48
C3	Wrappers	SVM	SVM: RBF kernel (cost(c), gamma (g)) grid search (5-fold CV) to find optimal c and g
C4	CFS	Naive Bayes	CFS: Greedy stepwise, Forward selection (No. of features to retain = 5)
C5	CFS	J48	CFS: Greedy stepwise, Forward selection (No. of features to retain = 5), Confidence Factor= 0.5 for J48
C6	CFS	SVM	CFS: Greedy stepwise, Forward selection (No. of features to retain = 5), SVM: RBF kernel (cost(c), gamma (g)) grid search (5-fold CV) to find optimal c and g

TABLE 3

ER-Positive vs. ER-Negative

Meta-classifier	PS = G-Q-m	Avg. features per fold	Avg. acc %	Sens. %	Spec. %
C1	256-32-17	4	84.44	80	89
C2	64-32-17	3	81.1	90	72
C3	64-32-17	3	76.11	80	72
C4	64-8-11	5	79.44	70	89
C5	64-8-13	5	84.44	80	89
C6	64-32-17	5	82.22	70	94

52.6% majority class, best results are in bold. Results are from initial enhancement kinetic map and rapid-delayed-washout habitat.

TABLE 4

FourOrMoreLN vs. NoLN

Meta-Classifer	PS=G-Q-m	Avg. features per fold	Avg. acc %	Sens. %	Spec. %
C1	64-8-11	3	82.29	83	81
C2	64-8-13	3	88.89	78	100
C3	128-8-13	3	76.04	83	69
C4	64-8-11	5	74	67	80
C5	64-8-13	5	88.89	78	100
C6	64-8-13	5	76.74	72	81

Slices chosen by first radiologist (52.9% majority class, best results are in bold). Results are from postinitial enhancement kinetic map and rapid-delayed-washout habitat.

TABLE 5

FourOrMoreLN vs. NoLN

Meta-Classifer	PS=G-Q-m	Avg. features per fold	Avg. acc %	Sens. %	Spec. %
C1	64-8-11	5	66.66	83.3	50
C2	128-8-13	3	82.9	72.2	93.75
C3	64-32-9	7	73.26	77.77	68.75
C4	64-8-11	5	69.79	83.3	56.25
C5	128-8-13	5	82.9	72.2	93.75
C6	64-16-9	5	73.26	77.77	68.75

Slices chosen by second radiologist (52.9% majority class, best results are in bold). Results are from postinitial enhancement kinetic map and rapid-delayed-washout habitat.

TABLE 6

Features Selected in More Than 65% of the Folds in Foldwise Feature Selection

Classification task based on	Meta-classifier	No. of features	Common features selected
ER status	Wrappers-NB	4	difference entropy, information measure of correlation, LBP-HF-22, LBP-HF-30
Lymph Node Status (1 st MD: J.D)	Wrappers-J48	3	information measure of correlation, LBP-HF-20, LBP-HF-24
Lymph Node Status (2 nd MD: R.W)	Wrappers-J48	6	information measure of correlation, LBP-HF-20, LBP-HF-21, LBP-HF-58, LBP-HF-19, LBP-HF-37