



**HAL**  
open science

# Comparison of Dynamic Contrast-Enhanced MRI and Non-Mono-Exponential Model-Based Diffusion-Weighted Imaging for the Prediction of Prognostic Biomarkers and Molecular Subtypes of Breast Cancer Based on Radiomics

Lan Zhang, Xin-xiang Zhou, Lu Liu, Ao-yu Liu, Wen-juan Zhao, Hong-xia Zhang, Yue-Min Zhu, Zi-xiang Kuai

► **To cite this version:**

Lan Zhang, Xin-xiang Zhou, Lu Liu, Ao-yu Liu, Wen-juan Zhao, et al.. Comparison of Dynamic Contrast-Enhanced MRI and Non-Mono-Exponential Model-Based Diffusion-Weighted Imaging for the Prediction of Prognostic Biomarkers and Molecular Subtypes of Breast Cancer Based on Radiomics. Journal of Magnetic Resonance Imaging, In press, 10.1002/jmri.28611 . hal-04210377

**HAL Id: hal-04210377**

**<https://hal.science/hal-04210377v1>**

Submitted on 19 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Comparison of Dynamic Contrast-Enhanced MRI and Non-Mono-Exponential Model-Based Diffusion-Weighted Imaging for the Prediction of Prognostic Biomarkers and Molecular Subtypes of Breast Cancer Based on Radiomics

L. Zhang, X. X. Zhou, L. Liu, A. Y. Liu, W. J. Zhao, H. X. Zhang, Y. M. Zhu, Z. X. Kuai

## Abstract

**Background:** Dynamic contrast-enhanced (DCE) MRI can be used to characterize breast cancer, but the side effects of gadolinium-based contrast agent limit its clinical application.

**Purpose:** To compare the performances of DCE-MRI, non-mono-exponential model-based diffusion-weighted imaging (NME-DWI) and their combination as multiparametric MRI (MP-MRI) in the prediction of breast cancer prognostic biomarkers and molecular subtypes based on radiomics.

**Study Type:** Prospective.

**Population:** 477 female patients (age  $51 \pm 8$  years) with 483 breast cancers.

**Field Strength/Sequence:** 3.0 T / NME-DWI with SE-EPI readout with monopolar diffusion gradients ( $b=0 \sim 2000$  s/mm<sup>2</sup>) and DCE-MRI with e-THRIVE readout (one pre-contrast and five post-contrast acquisitions).

**Assessment:** The tumor volumes were manually segmented by two radiologists, the radiomics features were extracted from the segmentations, and optimal feature subsets were selected using recursive feature elimination method. To predict the status of prognostic biomarkers (ER, PR, HER2 and Ki-67) and molecular subtypes (luminal A/B and triple negative) the following models were implemented: random forest, adaptive boosting, support vector machine, linear discriminant analysis, and logistic regression.

**Statistical Tests:** The model predictive performances were compared between the DCE-MRI, NME-DWI and MP-MRI datasets using the area under the receiver-operating characteristic curve (AUC) and the DeLong test.  $P < 0.05$  was considered significant.

**Results:** No significant differences ( $P=0.062 \sim 0.984$ ) were observed in the AUCs of models between the DCE-MRI (AUC=0.62~0.87) and NME-DWI (AUC=0.62~0.91) datasets with few exceptions, but the model performances on the two datasets were significantly poorer than on the MP-MRI dataset (AUC=0.68~0.93). Additionally, the random forest and adaptive boosting models (AUC=0.62~0.93) outperformed other three models (AUC=0.62~0.90).

**Data Conclusion:** NME-DWI was comparable with DCE-MRI in predictive power and could be used as an alternative technique for patients with gadolinium allergy history. Besides, MP-MRI achieved remarkable improvement in diagnostic accuracy than either DCE-MRI or NME-DWI.

## Keywords

breast cancer; radiomics; biomarkers; dynamic contrast-enhanced; diffusion-weighted; magnetic resonance imaging

## Introduction

Breast cancer is a highly heterogeneous disease with varying molecular subtypes which have shown differences in terms of their incidence, prognosis, treatment sensitivity, and recurrence-free and disease-specific survival<sup>1,2</sup>. Thus, it is necessary to achieve histopathological identification for cancer subtypes, in order to select the most appropriate therapy and predict the therapeutic response.

In clinical practice, the determination of breast cancer molecular subtypes prior to treatment is mainly based on the status (+/-) of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2), as well as the proliferation rate (Ki-67) from immunohistochemical (IHC) testing of biopsy samples<sup>3,4</sup>. In general, four intrinsic molecular subtypes are described: luminal A, luminal B, HER2 over-expression (HER2+), and triple negative (TN) subtypes<sup>5,6</sup>. Nevertheless, it should be noted that tissue biopsy is invasive and can capture only a localized snapshot of heterogeneous tumor, subject to sampling bias<sup>7</sup>. Therefore, there is a need for alternative methods to noninvasively identify prognostic biomarkers (receptor status and proliferation rate) and molecular subtypes of breast tumor in its entirety.

Recent studies showed that dynamic contrast-enhanced MRI (DCE- MRI) coupled with radiomic analyses yielded encouraging results regarding the exploration of breast cancer molecular mechanism. Leithner et al. reported diagnostic accuracies close to 80% for differentiating tumor phenotypes on breast DCE-MRI using radiomic method<sup>9</sup>. Lee et al. used radiomics features from DCE-MRI to identify receptor statuses and the area under curve values of classifiers reached 0.8<sup>10</sup>. On one hand, radiomics can implement the high-throughput extraction of quantitative features from medical images and subsequently make use of the valuable information to assist diagnosis<sup>11,12</sup>. On the other hand, DCE-MRI can not only offer insight into vascularization and perfusion, but also has higher spatial resolution and interobserver reproducibility compared to other imaging techniques such as mammography and ultrasound<sup>13,14</sup>. Nevertheless, it has to be emphasized that the side effects of gadolinium-based contrast agent used in DCE-MRI, such as nausea, low blood pressure, and renal function impairment, to some extent limit its clinical application<sup>15</sup>.

Given this, some unenhanced MRI techniques have gotten increasing study, and been tried to identify prognostic biomarkers and molecular subtypes. Relatedly, non-mono-exponential (NME) model-based diffusion-weighted imaging (DWI) has stood out from numerous techniques. The NME models mainly include the intravoxel incoherent motion (IVIM), diffusion kurtosis and stretched exponential models<sup>16-18</sup>. The parameters derived from these models can reflect tumor microstructure, spatial heterogeneity and microvascular perfusion<sup>19,20</sup>. In the literature, these model parameters have been demonstrated to be the potential indicators for the prediction of receptor statuses and proliferation rates<sup>21-23</sup>. To date, however, the radiomic analysis of NME-DWI for determining the prognostic biomarkers and molecular subtypes of breast cancer still needs to be explored further.

In this study, we aimed to explore whether NME-DWI is comparable with DCE-MRI in predicting breast cancer biomarkers and molecular subtypes based on radiomic assessment as an alternative for patients who have high allergic risks to gadolinium. Additionally, we aimed to investigate if the combination of both DCE-MRI and NME-DWI, as multiparametric MRI (MP-MRI) can further improve the predictive performance.

## Materials and methods

### *Patients*

This prospective study was approved by the institutional review board and written informed consent was obtained from all patients. Inclusion criteria were patients with suspicious breast cancer (detected by mammogram or ultrasound) who received breast MRI examinations from July 2018 to October 2021. Exclusion criteria were: preoperative

interventions and therapies, poor quality of images, with obvious artifacts, a maximum tumor diameter  $\leq 5$  mm, and incomplete histopathologic results about receptor status and proliferation rate. A total of 582 patients were enrolled in this study and 105 cases were excluded. For patients with multiple unilateral tumors, the largest tumor mass was selected. 6 patients had simultaneous bilateral cancers and each cancer was evaluated as a separate cancer. Finally, 477 patients with 483 lesions were included in this study (Fig. 1).

### ***MRI protocol***

MRI scan was performed using a 3.0T MRI system (Ingenia, Philips Healthcare, Eindhoven, The Netherlands) and a dedicated seven-channel bilateral breast coil. DWI was acquired before contrast administration using a single-shot spin-echo echo planar imaging sequence with monopolar diffusion-encoding gradients and inversion recovery fat suppression. The DWI readout parameters were: repetition time (TR)/echo time (TE)=6443/77 msec, flip angle=90°, field of view (FOV)=324×324 mm<sup>2</sup>, reconstruction matrix size=352×352, slice thickness/gap=5/1 mm, spatial resolution=0.92×0.92×5 mm<sup>3</sup>, and *b*-values=0, 10, 25, 50, 75, 100, 200, 400, 600, 800, 1000, 1500 and 2000 s/mm<sup>2</sup>. DCE images were acquired before (S0) and at five points (S1~S5) at 90 s intervals after the injection of a standard dose (0.1 mmol/kg) of gadolinium-based contrast agent using an enhanced T1 high resolution isotropic volume excitation sequence with fat suppression. The DCE readout parameters were: TR/TE =4.8/2.1 msec; flip angle=12°; FOV=350×350 mm<sup>2</sup>; reconstruction matrix size=784×784; slice thickness/gap=1/0 mm, and spatial resolution=0.45×0.45×1 mm<sup>3</sup>.

### ***Pathologic assessment***

For all patients, pathology results were obtained from surgical specimens and were reviewed for the ER, PR, and HER2 statuses, as well as the Ki-67 values. An Allred score from the IHC greater than 2 was considered positive for ER and PR<sup>6</sup>. For the determination of HER2 status, an IHC HER2 score of 3+, or a score of 2+ with an additional condition of HER2 gene amplification by the fluorescence in situ hybridization (FISH) was considered positive<sup>6</sup>. Positive Ki-67 was defined as expression  $\geq 14\%$ <sup>24</sup>. According to the Chinese Anti-Cancer Association Committee of the Breast Cancer Society guidelines for diagnosis and treatment of breast cancer, molecular subtypes were categorized into luminal A (ER and/or PR+, HER2-, Ki-67-), luminal B (ER and/or PR+, HER2-, Ki-67+; or ER and/or PR+, HER2+, and Ki-67±), HER2+ (ER- and PR-, HER2+), and TN (ER and PR-, HER2-) subtypes<sup>5</sup>. Additionally, tumor size, BI-RADS category, histological type and grade were recorded.

### ***Tumor segmentation***

Image registration was performed between dynamic frames and between *b*-value images for each tumor using an efficient subpixel image rigid transformation algorithm<sup>25</sup>. Then, the 3D tumor regions of DCE-MRI and NME-DWI were manually delineated by two radiologists (\*\* and \*\* with 7 and 15 years of experience in breast MRI) who were blinded to the pathologic results in a slice-by-slice manner on the dynamic frames and the *b*-value images that could best show the tumor boundaries relative to adjacent tissues, respectively<sup>25</sup>. Both radiologists reached consensus regarding all tumor segmentations.

### ***Image processing***

For each tumor in DCE-MRI, ten pharmacokinetic three-dimension (3D) parametric maps characterizing the physiological process of uptake and washout nature of contrast agent in the breast tumor during dynamic imaging series were calculated from the kinetic curves of enhancing voxels within the tumor. The ten 3D kinetic maps included: maximal uptake, time to peak, uptake rate, washout rate, curve shape index, enhancement at first postcontrast time point, percent enhancement, signal enhancement ratio, and early and late enhancement maps.

For each tumor in NME-DWI, the IVIM, diffusion kurtosis and stretched exponential 3D parametric maps were generated from the multi-*b*-value diffusion-weighted signals of voxels

within the breast tumor using a nonlinear least-squares fitting approach with bound constraints. The IVIM model parameters were computed on the basis of diffusion-weighted signals in the range of  $b=0\sim 800$  s/mm<sup>2</sup>, while the diffusion kurtosis and stretched exponential model parameters were obtained from diffusion-weighted signals of 0, 200, 800, 1000, 1500, 2000 s/mm<sup>2</sup>. The three models are expressed as follows:

1. IVIM model:

$$\frac{S(b)}{S_0} = (1-f)e^{-bD_s} + fe^{-bD_f} \quad (1)$$

where  $S_0$  represents  $S(b)$  for  $b=0$  s/mm<sup>2</sup> ( $b_0$ ),  $S(b)$  diffusion-weighted signal intensity at a given  $b$ -value,  $f$  the perfusion volume fraction,  $D_s$  the diffusion coefficient, and  $D_f$  the pseudo-diffusion coefficient<sup>16</sup>.

2. Diffusion kurtosis model:

$$\frac{S(b)}{S_0} = e^{(-bMD + \frac{1}{6}b^2MD^2MK)} \quad (2)$$

where  $MD$  stands for the mean diffusivity and  $MK$  the mean kurtosis expressing the deviation from Gaussian distribution<sup>17</sup>.

3. Stretched exponential model:

$$\frac{S(b)}{S_0} = e^{-(bDDC)^\alpha} \quad (3)$$

where  $DDC$  is the distributed diffusion coefficient,  $\alpha$  the anomalous exponent term characterizing the deviation from the mono-exponential decay ( $0 \leq \alpha \leq 1$ )<sup>18</sup>.  $\alpha=1$  represents homogeneous diffusion, while an  $\alpha=0$  represents highly heterogeneous diffusion<sup>18</sup>.

According to literature, the fitting boundaries of  $f$ ,  $D_s$ ,  $D_f$ ,  $MD$ ,  $MK$ ,  $DDC$  and  $\alpha$  were set as  $[0, 1]$ ,  $[0, 2] \times 10^{-3}$  mm<sup>2</sup>/s,  $[0, 60] \times 10^{-3}$  mm<sup>2</sup>/s,  $[0, 2] \times 10^{-3}$  mm<sup>2</sup>/s,  $[0, 2]$ ,  $[0, 2] \times 10^{-3}$  mm<sup>2</sup>/s and  $[0, 1]$ , respectively<sup>22,26,27</sup>. In addition, for avoiding local minima in fitting process, a multiple initialization strategy was adopted. The initial values of  $f$ ,  $D_s$ ,  $D_f$ ,  $MD$ ,  $MK$ ,  $DDC$  and  $\alpha$  parameters were set within their respective boundaries with a step size of 0.25,  $0.5 \times 10^{-3}$  mm<sup>2</sup>/s,  $10 \times 10^{-3}$  mm<sup>2</sup>/s,  $0.5 \times 10^{-3}$  mm<sup>2</sup>/s, 0.5,  $0.5 \times 10^{-3}$  mm<sup>2</sup>/s and 0.25, respectively.

The parametric maps derived from DCE-MRI and NME-DWI were given in Table 1.

### **Feature extraction and selection**

Before feature extraction, the following preprocessing steps were performed: normalized to  $\mu \pm 3\sigma$ , resampled to voxel size of  $1 \times 1 \times 1$  mm<sup>3</sup> using cubic interpolation, and discretized by a fixed bin width on the levels of grey, to eliminate the variance caused by different imaging sequences, avoid the anisotropic resolution, and improve the reproducibility<sup>28,29</sup>. Then, three categories of radiomics features were extracted from the original and derived 3D maps: morphological, histogram-based first-order, and gray level co-occurrence matrix (GLCM)-based second-order texture features (Table 2), using Pyradiomics package (v.3.0.1) in Python (v.3.9.12). Morphological features were calculated on the  $S_0$  maps of DCE-MRI and the  $b_0$  maps of NME-DWI, respectively. Histogram and GLCM features were computed on the kinetic maps from DCE-MRI and the model parametric maps from NME-DWI (Table 1). In total, for each tumor, 434 features and 308 features were extracted from the 3D maps from DCE-MRI and NME-DWI, respectively. For MP-MRI, the number of features was the sum of above two entries (i.e.,  $742=434+308$ ). Owing to the differences in units and magnitudes among radiomics features, the extracted features were subjected to z-score normalization by scaling values to a mean of 0 and a standard deviation of 1 prior to dimensionality reduction.

In the extracted features, some may be highly correlated to each other and less identifying capacity for prognostic biomarkers or molecular subtypes. Also, too many features could also increase the computational cost and the overfitting risk. Thus, feature selection is essential.

Recursive feature elimination (RFE) algorithm was utilized to obtain the optimal feature subset in view of their redundancy and distinguishing power<sup>30</sup>. The RFE feature selection was performed using the scikit-learn package (v.1.1.1) in Python (v.3.9.12). The feature subset with the best performance score was chosen for each classification task.

### ***Machine learning model-based classification***

In the present study, six classification tasks were accomplished successively: (1) ER+ vs. ER-, (2) PR+ vs. PR-, (3) HER2+ vs. HER2-, (4) Ki-67+ vs. Ki-67-, (5) luminal A/B vs. non-luminal A/B, and (6) TN vs. non-TN on the basis of the three imaging datasets (DCE-MRI, NME-DWI, and MP-MRI) using five machine learning models: random forest (RF), adaptive boosting (AB), support vector machine (SVM), linear discriminant analysis (LDA), and logistic regression (LR). Five-fold cross validation, in which each fold roughly had the same ratio for two classes, was used to assess the performance of the classifiers on each imaging dataset. In case of sample imbalance between two differentiated subgroups, which may adversely impact the performance of classifiers and lead to unreliable classification results, the synthetic minority oversampling technique (SMOTE) was applied to balance the datasets. By using SMOTE, the minority class was oversampled by taking each sample and introducing synthetic samples along the line segments joining any/all of the  $k$  (commonly  $k=5$ ) nearest neighbors<sup>31</sup>. Python scikit-learn package (v.1.1.1) and imbalanced package (v.0.9.1) were employed to implement classifiers and SMOTE technique, respectively.

The complete process of above radiomic analysis, including image processing, feature extraction and selection, as well as classification is exhibited in Figure 2.

### ***Statistical Analysis***

All statistical analyses were performed using scikit-learn package (v.1.1.1) in Python (v.3.9.12). The student's  $t$ -test was used to test intergroup differences in age and tumor size. The chi-square and Fisher's exact test were executed on the rest of the categorical characteristics to find whether the constituent ratio was significantly different between subgroups. For each classification task, the area under the receiver-operating characteristic (ROC) curve (AUC) values were calculated on all samples and were compared between models and between imaging datasets (DCE-MRI, NME-DWI, and MP-MRI) by the Delong test. Confidence intervals of AUC were obtained by using bootstrap analysis with 100 000-fold resampling<sup>32</sup>. The level of confidence was kept at 95% and results with  $P<0.05$  was considered to be statistically significant.

## **Results**

Clinical and pathologic characteristics of all patients are summarized in Table 3. No significant differences were found between subgroups for both of the patient ages ( $P=0.362\sim 0.725$ ) and lesion diameters ( $P=0.271\sim 0.834$ ). Likewise, no significant differences were found in the affected side between subgroups ( $P=0.103\sim 0.632$ ). But for other categorical characteristics, the significant difference can be observed in some cases, in particular the pathological type which was significantly different between subgroups in each classification task. Figure 3 shows the kinetic maps of DCE-MRI and the model parametric maps of NME-DWI from a 47-year-old female with invasive ductal cancer in the right breast.

The AUC values of five models for six classification tasks on three imaging datasets are provided in Table 4. Corresponding ROC curves are depicted in Figure 4. In terms of the AUC values obtained on different imaging datasets, it can be observed that there were no significant differences ( $P=0.062\sim 0.984$ ) between DCE-MRI (AUC=0.62~0.87) and NME-DWI (AUC=0.62~0.91) with few exceptions, but the performances of classifiers on MP-MRI (AUC=0.68~0.93) were significantly better than on both DCE-MRI and NME-DWI in many cases, as illustrated in Table 5.

As for the differences of performances between models (Figure 5), the RF model was

comparable with the AB model. Also, both of them (AUC=0.62~0.93) were superior to other three models (AUC=0.62~0.90) except in the differentiation of PR statuses. Concerning the SVM, LDA and LR models, the SVM model relatively outperformed the other two in classification performance, while there were no significant differences between the LDA and LR models in most cases.

## Discussion

This study demonstrated the potential of NME-DWI in predicting the prognostic biomarkers and molecular subtypes of breast cancer based on radiomics. NME-DWI was comparable with DCE-MRI in terms of predictive power without the need of injecting contrast agent. In addition, it's worth noting as well that the combination of DCE-MRI and NME-DWI, as MP-MRI, presented better classification performances than either imaging technique alone.

Although the application of radiomic-based NME-MRI in the exploration of breast cancer prognostic and molecular mechanisms is novel, DCE-MRI combined with radiomics has been widely used and assessed in related studies. Lee et al. obtained a median AUC of 0.80 for the differentiation of receptor statuses using the RF model on DCE images<sup>10</sup>. Agner et al. employed the SVM model to yield an AUC of 0.73 for distinguishing TN from non-TN<sup>34</sup>. In addition, according to Castaldo et al., the RF model achieved an AUC of 0.86 for the prediction of ER+ vs. ER-<sup>35</sup>. Overall, the diagnostic performances of DCE-MRI in this work are similar or superior to these previous results.

The most challenging and critical point in the radiomic analysis is tumor segmentation because the features data are obtained from the segmented volumes. In contrast with the diffusion-weighted images, the DCE-MRI images typically show preferable tissue contrast and shaper tumor outline due to higher spatial resolution and thus allows tumor margins to be delineated more accurately<sup>33</sup>. This may be the reason that DCE-MRI was superior to NME-DWI in individual cases. In this work, for minimizing selection biases, the tumor segmentations were performed by two radiologists in consensus.

Establishing a predictive model is crucial in radiomic analysis and suitable model selection could ensure the reliability and stability of the results<sup>10</sup>. In our study, the RF and AB models presented better predictive powers than other three models. Indeed, the RF and AB models are essentially ensemble classifiers consisting of a set of individually trained decision trees and can predict ensemble response by aggregating results from these weak learners<sup>36</sup>. Therefore, they allow the production of better diagnostic performance compared to a single model.

Our findings suggest three key advances with respect to radiomic assessment of breast cancer in 3.0T MRI. First, we performed a prospective study of 483 breast cancers, whereas most previous researches were retrospective and have relatively small sample sizes, which could cause the selection bias and affect the final results<sup>37</sup>. Second, we carried out full 3D volumetric analysis for each tumor instead of single-slice analysis which may miss important features because of intratumor heterogeneity. Third, the DWI sequence of this study covered a wide range of low, median, and high *b*-values. This means that more histological characteristics, such as tumor heterogeneity and capillary perfusion, can be derived and thus the predictive power of DWI could be fully showed.

### Limitations

First, we used immunohistochemical surrogates for breast cancer prognostic biomarkers instead of performing full genetic sequencing. Second, tumor segmentation in this work was not automatic, and thus could be subject to potential human error. Third, the classifier may have been biased by having multiple lesions from the same patient. Fourth, radiomic analysis was performed only once; hence, the inter- and intra-observer variabilities were not assessed. Finally, it is uncertain whether the present predictive models are applicable to variations in

imaging protocols and machines at other institutions.

#### Conclusion

NME-DWI achieved a comparable performance to DCE-MRI for predicting breast cancer prognostic biomarkers and molecular subtypes based on radiomics without the use of gadolinium contrast agents. Additionally, MP-MRI combining DCE-MRI and NME-DWI achieved remarkable improvement in predictive performance than either technique alone and could be recommended to maximize the diagnostic accuracy for patients without allergic history to gadolinium.

#### References

1. Martelotto LG, Ng CK, Piscuoglio S, Weigelt B, Reis-Filho JS. Breast cancer intra-tumor heterogeneity. *Breast Cancer Res.* 2014;16:210.
2. Nguyen PL, Taghian AG, Katz MS, et al. Breast cancer subtype approximated by estrogen receptor, progesterone receptor, and HER-2 is associated with local and distant recurrence after breast-conserving therapy. *J Clin Oncol.* 2008;26:2373-2378.
3. Gnant M, Harbeck N, Thomssen C. St. Gallen 2011: summary of the consensus discussion. *Breast care.* 2011;6:136-141.
4. Sutton EJ, Dashevsky BZ, Oh JH, et al. Breast cancer molecular subtype classifier that incorporates MRI features. *J Magn Reson Imaging.* 2016;44:122-129.
5. Association CA-C. Committee of breast cancer society guideline for diagnosis and treatment of breast cancer (Version 2019)(in Chinese). *China Oncol.* 2019;29:609-679.
6. Saha A, Harowicz MR, Grimm LJ, et al. A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features. *Br J Cancer.* 2018;119:508-516.
7. Luo HB, Du MY, Liu YY, et al. Differentiation between Luminal A and B Molecular Subtypes of Breast Cancer Using Pharmacokinetic Quantitative Parameters with Histogram and Texture Features on Preoperative Dynamic Contrast-Enhanced Magnetic Resonance Imaging. *Acad Radiol.* 2020;27:e35-e44.
8. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:4006.
9. Leithner D, Mayerhoefer ME, Martinez DF, et al. Non-invasive assessment of breast cancer molecular subtypes with multiparametric magnetic resonance imaging radiomics. *Journal of Clinical Medicine.* 2020;9:1853.
10. Lee JY, Lee K-s, Seo BK, et al. Radiomic machine learning for predicting prognostic biomarkers and molecular subtypes of breast cancer using tumor heterogeneity and angiogenesis properties on MRI. *Eur Radiol.* 2022;32:650-660.
11. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology.* 2013;278:563-577.
12. Grimm, Lars J. Breast MRI radiogenomics: Current status and research implications. *J Magn Reson Imaging.* 2016;43:1269-1278.
13. O'Connor JP, Jackson A, Parker GJ, Jayson GC. DCE-MRI biomarkers in the clinical evaluation of antiangiogenic and vascular disrupting agents. *Br J Cancer.* 2007;96:189-195.
14. Kuhl CK, Schrading S, Leutner CC, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *J Clin Oncol.* 2005;23:8469-8476.
15. Runge VM. Safety of the Gadolinium-Based Contrast Agents for Magnetic Resonance Imaging, Focusing in Part on Their Accumulation in the Brain and Especially the Dentate Nucleus. *Invest Radiol.* 2016;51:273-279.
16. Le BD, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology.* 1988;168:497-505.



17. Jensen JH, Helpert JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn Reson Med.* 2005;53:1432–1440.
18. Bennett KM, Schmainda KM, Bennett RT, Rowe DB, Lu H, Hyde JS. Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn Reson Med.* 2003;50:727-734.
19. Zhang XS, Liu EH, Wang XY, et al. Short-term repeatability of in vivo cardiac intravoxel incoherent motion tensor imaging in healthy human volunteers. *J Magn Reson Imaging.* 2021;55:854-865.
20. Zhou Y, Zhang H-X, Zhang X-S, et al. Non-mono-exponential diffusion models for assessing early response of liver metastases to chemotherapy in colorectal Cancer. *Cancer Imaging.* 2019;19:39.
21. You C, Li J, Zhi W, et al. The volumetric-tumour histogram-based analysis of intravoxel incoherent motion and non-Gaussian diffusion MRI: association with prognostic factors in HER2-positive breast cancer. *J Transl Med.* 2019;17:182.
22. Suo S, Cheng F, Cao M, et al. Multiparametric diffusion-weighted imaging in breast lesions: Association with pathologic diagnosis and prognostic factors. *J Magn Reson Imaging.* 2017;46:740-750.
23. Cho GY, Moy L, Kim SG, et al. Evaluation of breast cancer using intravoxel incoherent motion (IVIM) histogram analysis: comparison with malignant status, histological subtype, and molecular prognostic factors. *Eur Radiol.* 2016;26:2547-2558.
24. Veronese SM, Maisano C, Scibilia J. Comparative prognostic value of Ki-67 and MIB-1 proliferation indices in breast cancer. *Anticancer Res.* 1995;15:2717-2722.
25. Guizar-Sicairos M, Thurman ST, Fienup JR. Efficient subpixel image registration algorithms. *Opt Lett.* 2008;33:156-158.
26. Bedair R, Priest AN, Patterson AJ, et al. Assessment of early treatment response to neoadjuvant chemotherapy in breast cancer using non-mono-exponential diffusion models: a feasibility study comparing the baseline and mid-treatment MRI examinations. *Eur Radiol.* 2017;27:2726–2736.
27. Liu C, Wang K, Li X, et al. Breast lesion characterization using whole-lesion histogram analysis with stretched-exponential diffusion model. *J Magn Reson Imaging.* 2018;47:1701-1710.
28. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magn Reson Imaging.* 2004;22:81-91.
29. Mao B, Zhang LZ, Ning PG, Ding F, Ma JD. Preoperative prediction for pathological grade of hepatocellular carcinoma via machine learning-based radiomics. *Eur Radiol.* 2020;30:6924-6932.
30. Takahashi S, Takahashi W, Tanaka S, Haga A, Saito N. Radiomics Analysis for Glioma Malignancy Evaluation Using Diffusion Kurtosis and Tensor Imaging. *Int J Radiat Oncol Biol Phys.* 2019;105:784-791.
31. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of artificial intelligence research.* 2011;16:321-357.
32. Litjens G, Sánchez CI, Timofeeva N, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep.* 2016;6:26286.
33. Pinker K, Moy L, Sutton EJ, et al. Diffusion-weighted imaging with apparent diffusion coefficient mapping for breast cancer detection as a stand-alone parameter: comparison with dynamic contrast-enhanced and multiparametric magnetic resonance imaging. *Invest Radiol.* 2018;53:587-595.
34. Agner SC, Rosen MA, Englander S, 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*. 2014;272:91-99.

35. Castaldo R, Pane K, Nicolai E, Salvatore M, Franzese M. The impact of normalization approaches to automatically detect radiogenomic phenotypes characterizing breast cancer receptors status. *Cancers*. 2020;12:518.

36. Fields BK, Demirjian NL, Hwang DH, et al. Whole-tumor 3D volumetric MRI-based radiomics approach for distinguishing between benign and malignant soft tissue tumors. *Eur Radiol*. 2021;31:8522-8535.

37. Ye D-M, Wang H-T, Yu T. The application of radiomics in breast MRI: a review. *Technol Cancer Res Treat*. 2020;19:1533033820916191.

TABLE 1 The original and derived parametric maps from DCE-MRI and NME-DWI datasets and the corresponding features extracted from these maps.

Imaging	Model	Parametric map	Feature type
DCE-MRI	Original	$S_0$	Morphology
	Kinetics	Maximal uptake, time to peak, uptake rate, washout rate, curve shape index, enhancement at first postcontrast time point, PE, SER, $\text{map}_{\text{early enhance}}$ , $\text{map}_{\text{late}}$	Histogram and GLCM
NME-DWI	Original	$b_0$	Morphology
	IVIM	$f$ , $D_s$ , and $D_f$	Histogram and GLCM
	Diffusion kurtosis	$MD$ and $MK$	Histogram and GLCM
	Stretched exponential	$DDC$ and $\alpha$	Histogram and GLCM

PE: percent enhancement, which is defined as  $(I_{S1} - I_{S0})/I_{S0}$ ; SER: signal enhancement ratio, which is defined as  $(I_{S1} - I_{S0})/(I_{S5} - I_{S0})$ ;

$\text{map}_{\text{early enhance}}$ : early enhancement map, which is defined as  $I_{S1} - I_{S0}$ ;

$\text{map}_{\text{late enhance}}$ : late enhancement map, which is defined as  $I_{S5} - I_{S1}$ ;

IVIM: intravoxel incoherent motion;  $f$ : the perfusion volume fraction;  $D_s$ : the diffusion coefficient;  $D_f$ : the pseudo-diffusion coefficient;  $MD$ : mean diffusivity;

$MK$ : mean kurtosis;  $DDC$ : distributed diffusion coefficient;  $\alpha$ : the anomalous exponent term.

GLCM: gray-level co-occurrence matrices.

TABLE 2 Information of the extracted radiomics features.

Feature type	Radiomics features	No.
Morphology	Elongation, Flatness, Least Axis Length, Minor Axis Length, Maximum 2D diameter (Column), Maximum 2D diameter (Row), Maximum 2D diameter (Slice), Maximum 3D diameter, Mesh Volume, Major Axis Length, Sphericity, Surface Area, Surface Area to Volume ratio, Voxel Volume	14
Histogram	10 <sup>th</sup> Percentile, 90 <sup>th</sup> Percentile, Energy, Entropy, Interquartile Range, Kurtosis, Maximum, Mean Absolute Deviation (MAD), Mean, Median, Minimum, Range, Robust Mean Absolute Deviation (rMAD), Root Mean Squared (RMS), Skewness, Total Energy, Uniformity, Variance	18
GLCM	Autocorrelation, Cluster Prominence, Cluster Shade, Cluster Tendency, Contrast, Correlation, Difference Average, Difference Entropy, Difference Variance, Inverse Difference (ID), Inverse Difference Moment (IDM), Inverse Difference Moment Normalized (IDMN), Inverse Difference Normalized (IDN), Informational Measure of Correlation (IMC) 1, Informational Measure of Correlation (IMC) 2, Inverse Variance, Joint Average, Joint Energy, Joint Entropy, Maximal Correlation Coefficient (MCC), Maximum Probability, Sum Average, Sum Entropy, Sum of Squares	24

No.: number. GLCM: gray-level co-occurrence matrices.

TABLE 3 Clinical and pathologic characteristics of patient studies.

Characteristics	ER		PR		HER2		Ki-67		Luminal A/B		TN	
	P	N	P	N	P	N	≥14%	<14%	P	N	P	N
No. of patients	313	164	279	198	180	297	361	116	315	162	74	403
No. of lesions	318	165	282	201	180	303	364	119	320	163	75	408
Mean age (years)	52±8	49±9	50±10	51±9	48±11	53±9	51±13	50±11	52±5	49±7	54±9	47±8
BI-RADS												
IV	69	23	48	44	28	64	83	9	70	22	12	80
V	97	30	73	54	44	83	84	43	104	23	19	108
VI	152	112	161	103	108	156	197	67	146	118	44	220
Mean tumor diameter (mm)	32±21	33±22	32±19	34±21	33±23	34±22	34±21	33±22	32±22	33±21	34±23	32±21
Pathological type												
Invasive ductal cancer	217	92	176	133	83	226	236	73	212	97	50	259
Invasive lobular cancer	63	40	73	30	62	41	82	21	69	34	19	84
Mixed invasive ductal and lobular cancer	26	25	22	29	25	26	32	19	24	27	0	51
Mucinous cancer	4	0	4	0	0	4	1	3	4	0	0	4
Others	8	8	7	9	10	6	13	3	11	5	6	10
Grade												
I	53	46	53	46	49	50	71	28	51	48	15	84
II	183	89	133	139	95	177	198	74	194	78	40	244
III	82	30	96	16	36	76	95	17	75	37	20	80
Affected side												
Right	166	82	130	104	93	142	176	61	147	86	43	191
Left	152	83	152	97	87	161	188	58	173	77	32	217

Data are numbers of lesions unless otherwise indicated. Others include invasive papillary carcinoma, invasive micropapillary carcinoma and metaplastic carcinoma. P: positive; N: negative; No.: number; ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2; Luminal A/B: Luminal A and Luminal B; TN: triple negative; BI-RADS: breast imaging-reporting and data system.

TABLE 4 AUCs of five machine learning models for six classification tasks on three imaging datasets.

Classification task	Classifier	DCE-MRI	NME-DWI	MP-MRI
ER+ vs. ER-	RF	0.86 [0.81, 0.87]	0.87 [0.85, 0.89]	0.90 [0.86, 0.92]
	AB	0.83 [0.80, 0.86]	0.84 [0.80, 0.86]	0.89 [0.87, 0.92]
	SVM	0.72 [0.68, 0.76]	0.77 [0.73, 0.80]	0.81 [0.78, 0.84]
	LDA	0.67 [0.64, 0.72]	0.67 [0.63, 0.70]	0.74 [0.69, 0.77]
	LR	0.68 [0.64, 0.72]	0.69 [0.64, 0.73]	0.73 [0.68, 0.76]
PR+ vs. PR-	RF	0.72 [0.66, 0.76]	0.71 [0.66, 0.76]	0.73 [0.65, 0.75]
	AB	0.68 [0.63, 0.73]	0.62 [0.57, 0.67]	0.71 [0.65, 0.75]
	SVM	0.70 [0.63, 0.73]	0.70 [0.64, 0.73]	0.72 [0.66, 0.76]
	LDA	0.66 [0.59, 0.69]	0.65 [0.61, 0.68]	0.70 [0.64, 0.71]
	LR	0.65 [0.59, 0.69]	0.68 [0.62, 0.72]	0.71 [0.65, 0.75]
HER2+ vs. HER2-	RF	0.85 [0.81, 0.88]	0.87 [0.86, 0.89]	0.88 [0.85, 0.90]
	AB	0.84 [0.81, 0.87]	0.83 [0.79, 0.85]	0.88 [0.85, 0.91]
	SVM	0.70 [0.66, 0.74]	0.71 [0.65, 0.73]	0.77 [0.74, 0.81]
	LDA	0.64 [0.61, 0.70]	0.69 [0.63, 0.71]	0.72 [0.67, 0.75]
	LR	0.64 [0.59, 0.68]	0.68 [0.63, 0.71]	0.70 [0.66, 0.74]
Ki-67+ vs. Ki-67-	RF	0.83 [0.81, 0.87]	0.83 [0.81, 0.88]	0.87 [0.84, 0.89]
	AB	0.83 [0.79, 0.87]	0.91 [0.88, 0.93]	0.93 [0.89, 0.95]
	SVM	0.73 [0.66, 0.75]	0.74 [0.68, 0.77]	0.85 [0.81, 0.88]
	LDA	0.62 [0.59, 0.67]	0.66 [0.63, 0.69]	0.68 [0.64, 0.71]
	LR	0.65 [0.61, 0.70]	0.64 [0.59, 0.68]	0.71 [0.67, 0.75]
Luminal A/B vs. non-luminal A/B	RF	0.86 [0.83, 0.88]	0.86 [0.82, 0.88]	0.88 [0.84, 0.89]
	AB	0.85 [0.81, 0.88]	0.85 [0.82, 0.88]	0.87 [0.84, 0.89]
	SVM	0.74 [0.70, 0.78]	0.76 [0.72, 0.79]	0.80 [0.76, 0.83]
	LDA	0.66 [0.62, 0.69]	0.68 [0.66, 0.74]	0.74 [0.68, 0.76]
	LR	0.66 [0.62, 0.70]	0.69 [0.64, 0.72]	0.74 [0.70, 0.78]
TN vs. non-TN	RF	0.86 [0.82, 0.87]	0.85 [0.82, 0.86]	0.90 [0.87, 0.92]
	AB	0.87 [0.84, 0.91]	0.88 [0.85, 0.89]	0.89 [0.87, 0.92]
	SVM	0.84 [0.80, 0.87]	0.79 [0.75, 0.82]	0.90 [0.86, 0.92]
	LDA	0.70 [0.67, 0.73]	0.73 [0.70, 0.77]	0.75 [0.72, 0.79]
	LR	0.62 [0.57, 0.67]	0.71 [0.67, 0.75]	0.75 [0.71, 0.79]

Numbers in square brackets are 95% confidence intervals (CIs). ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2; Ki-67+: Ki-67 $\geq$ 14%; Ki-67-: Ki-67<14%; Luminal A/B: Luminal A and Luminal B; TN: triple negative; RF: random forest; AB: adaptive boosting; SVM: support vector machine; LDA: linear discriminant analysis; LR: logistic regression.

TABLE 5 Comparisons ( $P$  values) of the performances of classifiers between different imaging datasets.

Classification task	Classifier	DCE-MRI	MP-MRI	MP-MRI
		vs. NME-DWI	vs. DCE-MRI	vs. NME-DWI
ER+ vs. ER-	RF	0.498 (-)	0.007 (↑↑)	0.001(↑↑)
	AB	0.984 (-)	0.003 (↑↑)	0.004 (↑↑)
	SVM	0.025 (↓)	0.001 (↑↑)	0.079 (-)
	LDA	0.799 (-)	0.024 (↑)	0.023 (↑)
	LR	0.926 (-)	0.034 (↑)	0.223 (-)
PR+ vs. PR-	RF	0.968 (-)	0.731 (-)	0.762 (-)
	AB	0.035 (↑)	0.674 (-)	0.041 (↑)
	SVM	0.361 (-)	0.861 (-)	0.455 (-)
	LDA	0.104 (-)	0.579 (-)	0.043 (↑)
	LR	0.968 (-)	0.017 (↑)	0.762 (-)
HER2+ vs. HER2-	RF	0.387 (-)	0.149 (-)	0.022 (↑)
	AB	0.421 (-)	0.041 (↑)	0.004 (↑↑)
	SVM	0.881 (-)	0.005 (↑↑)	0.002 (↑↑)
	LDA	0.042 (↓)	0.003 (↑↑)	0.369 (-)
	LR	0.236 (-)	0.023 (↑)	0.271 (-)
Ki-67+ vs. Ki-67-	RF	0.043 (-)	<0.001 (↑↑)	<0.001 (↑↑)
	AB	0.001 (↓↓)	<0.001 (↑↑)	<0.001 (↑↑)
	SVM	0.586 (-)	0.036 (↑)	0.031 (↑)
	LDA	0.382 (-)	0.002 (↑↑)	0.047 (↑)
	LR	0.635 (-)	0.024 (↑)	0.021 (↑)
Luminal A/B vs. non-luminal A/B	RF	0.426 (-)	0.895 (-)	0.486 (-)
	AB	0.860 (-)	0.374 (-)	0.465 (-)
	SVM	0.460 (-)	0.037 (↑)	0.170 (-)
	LDA	0.041 (↓)	0.010 (↑↑)	0.037 (↑)
	LR	0.429 (-)	0.004 (↑↑)	0.044 (↑)
TN vs. non-TN	RF	0.564 (-)	0.063 (-)	0.013 (↑↑)
	AB	0.820 (-)	0.633 (-)	0.769 (-)
	SVM	0.037 (↑)	0.007 (↑↑)	<0.001 (↑↑)
	LDA	0.062 (-)	0.038 (↑)	0.330 (-)
	LR	0.004 (↓↓)	<0.001 (↑↑)	0.137 (-)

↑ (or ↓) indicates a significant increase (or decrease) of AUC on the former imaging dataset compared to the latter one by DeLong test. ↑↑ (or ↓↓) indicates a significant level of  $P<0.01$ . (-) indicates no significant difference. ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2; Ki-67+: Ki-67 $\geq$ 14%; Ki-67-: Ki-67<14%; Luminal A/B: Luminal A and Luminal B; TN: triple negative; RF: random forest; AB: adaptive boosting; SVM: support vector machine; LDA: linear discriminant analysis; LR: logistic regression.

FIGURE 1 Flowchart of study population with inclusion and exclusion criteria. ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2.

FIGURE 2 Summary of radiomic analysis performed in this study. From left to right, image processing, feature extraction and selection and classification. S0: precontrast; IVIM: intravoxel incoherent motion; DKI: diffusion kurtosis imaging; SEI: stretched exponential imaging; RFE: recursive feature elimination; RF: random forest; AB: adaptive boosting; SVM: support vector machine; LDA: linear discriminant analysis; LR: linear regression; GLCM: gray-level co-occurrence matrices.

FIGURE 3 A case example from a 47-year-old female with invasive ductal cancer in the right breast. (**a~n**) DCE images and derived parametric maps: **a** pre-contrast image; **b** post-contrast image; **c** segmented ROI image; **d** 3D tumor outline image; **e** maximal uptake map, **f** time to peak map, **g** uptake rate map, **h** washout rate map, **i** curve shape index map, **j** enhancement map at first postcontrast time point; **k** percent enhancement; **l** signal enhancement ratio map; **m** enhancement map at early postcontrast phase; **n** enhancement map at late postcontrast phase. These color-coded parametric maps derived from DCE images were normalized on the basis of their respective ranges of parameter values. (**o~y**) DWI and derived parametric maps: **o** unweighted image; **p** DWI of  $b = 200 \text{ s/mm}^2$ ; **q** segmented ROI image; **r** 3D tumor outline image; **s**  $f$  map; **t**  $D_s$  map; **u**  $D_f$  map; **v**  $MD$  map; **w**  $MK$  map; **x**  $DDC$  map; **y**  $\alpha$  map. These color-coded parametric maps derived from DWI were normalized on the basis of their respective fitting boundaries.

FIGURE 4 The ROC curves of five machine learning models for six classification tasks on three imaging datasets (DCE-MRI, NME-DWI and MP-MRI). Numbers in parentheses are AUCs.

FIGURE 5 Comparisons of the performances between different classifiers. ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2; Lum A/B: Luminal A and Luminal B; TN: triple negative; RF: random forest; AB: adaptive boosting; SVM: support vector machine; LDA: linear discriminant analysis; LR: linear regression.