Radiogenomic signatures of the Oncotype DX recurrence score enhance the prediction of survival and treatment response in breast cancer: A multicohort study

Ming Fan[1], Yajing Cui[1], Chao You[2], Li Liu[2], Yajia Gu[2], Weijun Peng[2], Qianming Bai[2], Xin Gao[3], Lihua Li[1*]

[1]Institute of Biomedical Engineering and Instrumentation, Hangzhou Dianzi University, Hangzhou 310018, China

[2]Department of Radiology, Fudan University Shanghai Cancer Center, Shanghai, China

[3]Computational Bioscience Research Center (CBRC), Computer, Electrical and Mathematical Sciences and Engineering Division (CEMSE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

Correspondence:

Lihua Li,

1Institute of Biomedical Engineering and Instrumentation, Hangzhou Dianzi University, Hangzhou 310018, China

Phone: +86-571-86878587
Fax: +86-571-87713528
Email: lilh@hdu.edu.cn

Running title: Radiogenomic signatures of Oncotype DX recurrence score

Acknowledgment

This work was supported in part by grants from the National Natural Science Foundation of China: 61731008 (to L. Li) and 61871428 (to M. Fan) and the Natural Science Foundation of Zhejiang Province of China: LJ19H180001 (to M. Fan).

Disclosure of Potential Conflicts of Interest
The authors declare no potential conflicts of interest.

Word count: 4556
Total number of figures and tables: 6

Translational Relevance
The oncotype DX recurrence score (RS) assay serves as a reliable predictor of prognosis/recurrence and in decision making for neoadjuvant therapy in early-stage breast cancer. Radiomics provides intrinsic characteristics within heterogeneous tumors for more accurate modeling of the Oncotype DX RS. The clinical implications of the radiomics-derived prognostic signatures in treatment and prognosis have never been independently validated. Notably, we prosed a multicohort analysis strategy to develop radiogenomic signatures associated with Oncotype RS and to validate their effectiveness in treatment and survival analysis. Specifically, the radiogenomic signatures derived from genomic assays were validated as independent predictors for response to treatment and prognosis (recurrence-free survival and overall survival). The prediction performance was also augmented by combining the identified signatures with complementary imaging features. This study provides a noninvasive approach to identify biomarkers that could be promising for use in the diagnosis and prognosis of breast cancer.

Abstract

Purpose: Gene expression-determined prognostic biomarkers distinguish patients with different prognoses or responses to treatment. The purpose of this study was to identify preoperative radiogenomic signatures of breast cancer heterogeneity associated with the genomic Oncotype DX recurrence score (RS) and to validate their effectiveness with respect to prognosis and treatment outcomes in patients from multiple cohorts.

Experimental Design: In this retrospective study, three data cohorts were analyzed. The radiogenomic development cohort (n=130) was used to identify radiogenomic signatures by associating dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) with the matched Oncotype RS data. The prognostic implication of the imaging signatures was assessed by both recurrence-free survival (RFS) and overall survival (OS) in the prognostic validation cohort (n=88) collected from The Cancer Imaging Archive (TCIA). The therapeutic implication of the radiogenomic signatures was evaluated by predicting response to neoadjuvant chemotherapy (NACT) in the treatment validation cohort (n=114).

Results: Radiogenomic signatures were identified to generate the predicted Oncotype RS (pRS) (R square 0.332). High pRS greater than 29.9 was significantly associated with both RFS and OS (p=0.0012 and 0.0014, respectively), and the association results remained significant after controlling for histological information. A high pRS was significantly correlated with a good response to NACT (p=0.001). Multivariate analysis by combining the pRS and complementary imaging features achieved improved performance in both the prediction of response to NACT and survival.

Conclusions: Radiogenomic signatures associated with genomic assays were identified and validated across multiple cohorts, which indicate promising, noninvasive biomarkers of prognosis and treatment in breast cancer.

Keywords: Breast cancer; Neoadjuvant chemotherapy; Radiogenomic; Dynamic contrast-enhanced magnetic resonance imaging; Prognosis.

Introduction

Breast cancer is the most common female neoplasm. Neoadjuvant chemotherapy (NACT) is increasingly being utilized as the standard therapy for the management of patients with high-risk, localized, unresectable tumors [1]. Additionally, NACT can expedite the development and approval of treatments for certain patients, including those with the aim of undergoing breast-conserving surgery [2]. A comprehensive meta-analysis showed that patients who achieved pathologic complete response (pCR) after NACT were associated with improved overall survival (OS) or event-free survival (EFS) [3].

Molecular biomarkers such as estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and Ki-67 expression are commonly used for prognosis analysis and patient treatment response prediction [4]. To date, multigene molecular tests [5-8] have become available and provide early indicators of prognosis. Among these, Oncotype DX is a validated multigene (21 genes) recurrence score (RS) assay that serves as a reliable predictor of prognosis, i.e., recurrence, survival [9], treatment, and response to NACT [10-12]. Oncotype DX can impact decision making for surgery [13] and adjuvant therapy [14, 15] in early-stage breast cancer. The clinical application of Oncotype DX RS has provided improved survival [16, 17] and decreased rates of receiving adjuvant chemotherapy in breast cancer patients [18]. However, the high cost and invasive nature of the 21-gene assay may impede its more extensive applications in clinical practice [19, 20]. Previous studies have used easily available clinicopathologic parameter-based predictive models as surrogates for Oncotype RS [21]. However, these parameters, also acquired using invasive biopsy, remain limited in modeling information-rich genomic assays. Additional technology providing intrinsic characteristics within heterogeneous tumors is indispensable for more accurate prediction of the Oncotype DX RS. The development of noninvasive radiogenomic signatures from genomic assays is needed for improved clinical decision making for breast cancer patients.

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is widely used in breast cancer management, as it provides noninvasive, volumetric characterization of tumor morphologic and functional information. Radiomic analyses of tumor heterogeneity were conducted to correlate quantitative DCE-MRI features with Oncotype DX RS in early-stage, ER-positive invasive breast cancers [22-25]. As shown in a previous study, texture features derived from tumors are associated with prognosis [26] and treatment responses in breast cancer [27]. Existing studies have identified imaging signatures that were purposely designed based on their associations with the RS in breast cancer, but their clinical implications have never been validated. Whether the prognostic signatures could be independently used or augment the prediction of treatment outcomes or the survival/recurrence status remains unclear.

The purpose of this study was to identify radiogenomic signatures of breast tumor heterogeneity associated with Oncotype DX RS and, more importantly, to validate the clinical implications of the signatures in predicting the treatment response to NACT and the survival of breast cancer patients based on a multicohort study. Our study aimed not only to identify prognostic/therapeutic radiogenomic signatures but also to use them as independent and additional biomarkers for treatment decision making and prognostic assessment in breast cancer.

Methods

Overview of the study design and cohorts

Our analysis was performed on three independent cohorts: a radiogenomic development cohort, a prognostic validation cohort and a treatment validation cohort (all designed specifically for the purpose of this study). A three-stage study was performed as illustrated in Figure 1. First, radiogenomic signatures were identified based on their associations with the Oncotype DX RS in the radiogenomic development cohort. A regression model with the identified signatures was established and applied in the independent prognostic validation cohort, where the predicted RS (pRS) was generated. The prognostic implications of pRS were validated for their associations with the OS and recurrence-free survival (RFS) of breast cancer patients. pRS was also produced in the treatment validation cohort, where the prognostic implication was validated by predicting the response to NACT. The predictive power of pRS was augmented in a multivariate model by incorporating other additional and complementary radiomic features to enhance the prediction of the response to NACT and survival outcomes. The clinical characteristics of the patients are summarized in Supplementary Table 1. No significant difference regarding the histological characteristics was found between the groups. A detailed description of the three cohorts is illustrated in the following sections.

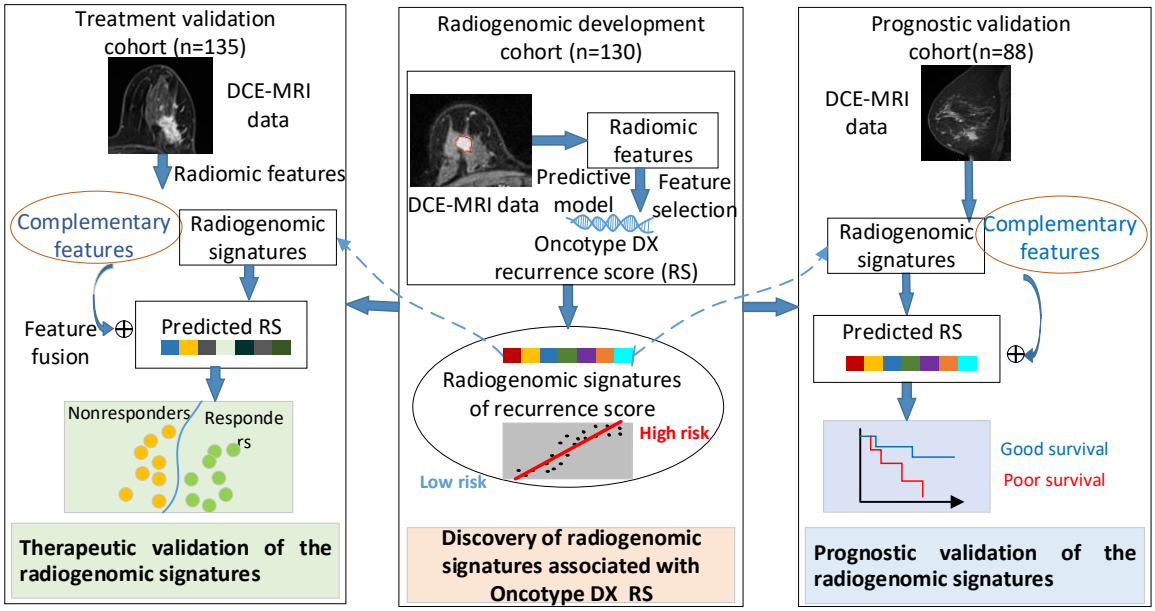

Figure 1. Framework overview

Radiogenomic development cohort: study population and imaging protocol

For the radiogenomic development cohort, breast cancer patient data were collected from Fudan Cancer Hospital following approval from the institutional review board (IRB). Informed consent was waived due to the retrospective nature of this study. This data cohort initially included 248 patients. Patients (n=118) were removed from this cohort for the following reasons: 1) no available image data or incomplete image data (n=20); 2) no pathology report (n=2); 3) surgery was performed before images were acquired (n=51); 4) no visible tumor or a diffuse tumor (n=35); and 5) low-quality genetic assays (n=10). The final cohort included 130 patients with imaging and RS data (Supplementary Table 1). All patients were hormone receptor (HR)-positive (ER- or PR-positive). Among these patients, 12 (9.2%) had a low score (RS<18), 72 (55.4%) had an intermediate score (18≤RS≤30) and 46 (35.4%) had a high score (RS>30).

DCE-MRI was performed using a dedicated 1.5-T breast MRI system in the prone position (Aurora Dedicated Breast MRI Systems, USA). For each patient, one precontrast image and three postcontrast image sequences were acquired at 120, 245, and 371 seconds after beginning the intravenous administration of gadobutrol. The following imaging parameters were used: repetition time (TR) = 29 ms, echo time (TE) = 4.8 ms, flip angle (FA) = 90°, field of view (FOV) = 360×360 mm, acquisition matrix =512×512 mm, slice thickness = 1.482 mm, and in-plane resolution = 0.7 mm.

Treatment validation cohort: study population and imaging protocol

For the treatment validation cohort, breast cancer patient data were collected from Fudan Cancer Hospital following approval from the IRB. This cohort initially included 442

patients with preoperative DCE-MRI data and matched clinicopathologic information. After excluding 307 patients who did not meet the selection criteria, 135 remained for analysis. Patients were excluded for the following reasons: 1) surgery was performed before MRI (n=43); 2) no available pathology report (n=8); 3) negative for ER or PR (n=146); and 4) no visible tumor or a diffuse tumor (n=49). Among these patients, 97 (71.9%) were nonresponders, and 38 (28.1%) were responders.

DCE-MRI was performed in the prone position using a Skyra scanning system (Siemens Skyra 3.0 T scanner, Germany). The following parameters were used: TR=4.5 ms, TE=1.56 ms, FA=10°, FOV=360×360 mm, matrix = 384×384 mm, slice thickness=2.2 mm, and in-plane resolution = 0.9375 mm. The time interval between S1 and S0 was 1 minute and 34 seconds, and the time interval between postcontrast slices was 43 seconds.

Prognostic validation cohort: study population and imaging protocol

The prognostic validation cohort included imaging and matched prognostic data collected from the I-SPY1 cohort in The Cancer Imaging Archive (TCIA) (http://www.cancerimagingarchive.net/). Patients with DCE-MRI and survival data, including RFS and OS, were analyzed. The cohort initially included 222 patients, and 143 were excluded for the following reasons: 1) no visible tumor or a diffuse tumor (n=24); 2) ER- or PR-negative (n=96); no image or an incomplete sequence (n=17); and 3) no histology report (n=1). This generated 88 patients for analysis. There were 17 (19.3%) recurrences after a mean follow-up period of 47.109 months (range, 7.200-79.567 months). The mean time to recurrence was 30.039 months (range, 7.200-55.967 months). The mean time to death was 29.071 months (range 10.567-55.967 months).

Imaging was performed using the following protocols: TR≤20 ms, TE = 4.5 ms, flip angle ≤45°, FOV 16-18 mm, and matrix 256×192. For each image, 64 slices with a thickness ≤2.5 mm and an in-plane resolution ≤1 mm were used. After the injection of contrast agent with a time interval of 2 minutes and 15 seconds, the first postcontrast image series was acquired. The second and third postcontrast images were acquired with time intervals of 7 minutes and 15 seconds and 7 minutes and 45 seconds, respectively.

Pathologic analysis

For the radiogenomic cohort, Oncotype DX RS was determined based on preoperative biopsy [8, 28] and was categorized into three levels: low (<18), intermediate (18-30), and high (≥31) [8]. For the treatment validation cohort, the response to NACT was evaluated with the surgical specimen-determined Miller Payne (MP) grading system [29]. Tumor response was categorized into five classes using MP scores (ranging from 1 to 5). As suggested in a previous study, patients with a tumor grade of 4 or 5 (total cell loss of more than 90%; categorized as almost pCR or pCR, respectively) were defined as responders [30]. The other patients with a tumor grade of 1, 2 or 3 (total cell loss of up to 90%) were defined as nonresponders. ER, PR and Ki-67 statuses were determined according to immunohistochemistry (IHC) with streptavidin-peroxidase (SP) detection [31, 32]. Ki-67 positivity was defined as an expression level greater than 14%. HER2

positivity was defined as IHC 3+ or 2+ with confirmation of gene amplification by fluorescence in situ hybridization (FISH)[32]. HR positivity was determined as ER and/or PR positivity.

Image preprocessing

Tumors were semiautomatically segmented following a previously reported procedure [33]. The center of the tumor was first manually annotated by a radiologist with 10 years of experience. Next, tumor areas were segmented automatically using the spatial fuzzy C-means (FCM) method [34]. Image normalization was performed to alleviate bias caused by varied imaging parameters in different data cohorts used in this study. First, all the images were resized to ensure that they had the same spatial resolution (0.703125×0.703125×1.48148 mm). Furthermore, the pixel value for the images in the three cohorts was mapped to a fixed gray area ranging from 0 to 800.

Radiomic analysis

Radiomic analysis was performed using the publicly available software Pyradiomics [35]. Tumor imaging features (n=107) included size/morphology (n=14), statistics (n=18), and texture (n=75) features. The texture features were calculated from the gray-level cooccurrence matrix (GLCM) (n=24), gray-level size zone matrix (GLSZM) (n=16), gray-level run length matrix (GLRLM) (n=16), gray-level difference matrix (n=14) and neighborhood gray tone difference matrix (NGTDM) (n=5). The precontrast image (S0), the intermediate postcontrast image (SM), which usually represents the highest enhancement value, and the postcontrast image series (SL), which represents a washout two and four minutes after the acquisition of S0, were analyzed. Features were calculated on the S0 images and the subtraction images between SM and S0 and between SL and S0, which were termed SM0 and SL0, respectively. Additionally, radiomic features were calculated on enhancement rate maps formulated as follows: (SM-S0)/S0 and (SL-SM)/(SM-S0) (termed M1 and M2, respectively). Therefore, a total of 479 imaging features were obtained from the DCE-MR image series and enhancement rate maps. Supplementary Table 2 shows the calculated features.

Statistical analysis

For the radiogenomic cohort, the radiogenomic signature was established using an elastic net regression model by combining the lasso (L_1-norm) and ridged (L_2-norm) penalizations [36]. To reduce multicollinearity of the features in the model, features with high similarity were removed. Specifically, for feature pairs with a Pearson correlation coefficient (PCC) greater than 0.9, the feature that had a higher average PCC than the other features were removed. The regularization parameters λ and α (for the L_1- and L_2-norms, respectively) were tuned by using a grid search under 10-fold cross validation (CV). Feature selection was performed using recursive feature elimination (RFE) to generate an optimal feature subset for prediction. The predictive model with the radiogenomic signatures and the tuned model parameters was established using all the samples in the radiognomic development cohort. The model was applied to the

prognostic validation and treatment validation cohorts to generate pRS, which was validated by treatment response and survival outcome prediction.

For the prognostic validation cohort, survival analysis was performed using Kaplan-Meier curves. Patients alive without an event were censored at 10 years. The Cox proportional hazards model was used to investigate whether the pRS added independent information in the presence of the clinical covariates of age, histological status, and tumor size. The optimal threshold of the pRS was determined by separating the patients with the most significant differences in survival. The likelihood ratio (LR) test was used to evaluate the significance of imaging features on survival.

For the treatment validation cohort, a two-sample t-test was performed to compare the difference in pRS between the responders and the nonresponders. A support vector machine (SVM) was utilized to discriminate the nonresponders from the responders. The area under the receiver operating characteristic (ROC) curve (AUC) was calculated to assess the predictive performance of the model in classification. Comparison of model performance in terms of AUC values was performed using the bootstrap method using the pROC package in the R program.

The false discovery rate was controlled using the Benjamini and Hochberg method. A corrected two-tailed p value less than 0.1 was considered significant. Statistical analysis was performed using the R program (version 4.0; R Foundation for Statistical Computing, Vienna, Austria). Machine learning methods of regression and classification were performed using Python. Image processing and tumor segmentation were performed using MATLAB (R2018a).

Results

Discovery of radiogenomic signatures based on their associations with Oncotype DX RS

In the first stage, radiogennomic signatures were identified in the radiognomic development cohort. Based on 10-fold CV, the optimal subset of 11 features in the multivariate regression model was used to examine its association with the Oncotype DX RS (Table 1 and Supplementary Figure 1f). The optimized parameters in the elastic net model were as follows: λ=2.68e-5 and α=0.131. The established elastic net achieved an R squared value of 0.332 (adjusted R squared value of 0.270), with p=8.29e-7. A Pearson correlation map of the 11 features is shown in Supplementary Figure 2. The signatures consist of morphologic and first-order features and texture features (GLCM-, GLSZM-, and GLRLM-based features).

Table 1. Multivariate regression model for associations between radiomic features and the Oncotype recurrence score

Feature type	Feature name	Coefficient	Correlation	p
Morphologic				
	Sphericity[1]	-2.103	0.45	5.89e-8
	Flatness[1]	-1.178	0.36	2.51e-5
First order				
	Skewness[1]	1.872	0.27	0.002
Texture				
GLCM				
	MCC[3]	-0.777	0.32	2.55e-4
	Joint Average[1]	0.510	0.26	0.003
	Difference Average[5]	2.093	0.23	0.008
GLSZM				
	Zone Entropy[5]	0.694	0.39	3.63e-6
	Zone Entropy[2]	-0.434	0.29	7.77e-4
GLRLM				
	Run Entropy[3]	1.833	0.30	5.06e-4
	Run Entropy[4]	-0.984	0.28	0.001
	Run Entropy[5]	2.352	0.27	0.002

[1]S0; [2]SM0; [3]SL0; [4]M1; [5]M2; MCC= maximal correlation coefficient; GLSZM=gray-level size zone matrix GLCM=gray-level cooccurrence matrix; GLRLM=gray-level run length matrix.

Individual imaging signatures were investigated to assess their correlations with the Oncotype RS. Examples of patients and RS signatures for the low- and high-risk groups are shown in Figure 2. Tumors with lower RS values showed higher levels of sphericity (Figure 2a-c), a measure of the roundness of the tumor region relative to a sphere, where a value of 1 indicates a perfect sphere. On the other hand, tumors with a higher RS showed a lower level of zone entropy, a heterogeneity-related texture feature that measures uncertainty/randomness in the distribution of zone sizes and gray levels (Figure 2d-f). Scatter plots of the morphologic, statistical and texture features are illustrated in Supplementary Figure 1a-e. Among these, the skewness (first-order feature) obtained in the precontrast image showed a positive correlation with the RS, i.e., negative skewness of tumor pixel value distribution likely represents a low RS (Supplementary Figure 1c). This result indicates that the presence of voxel values toward the lower values (left tail on a distribution) is associated with a low risk of recurrence. In this stage, the radiogenomic signatures were established by associating the imaging features with the Oncotype RS.

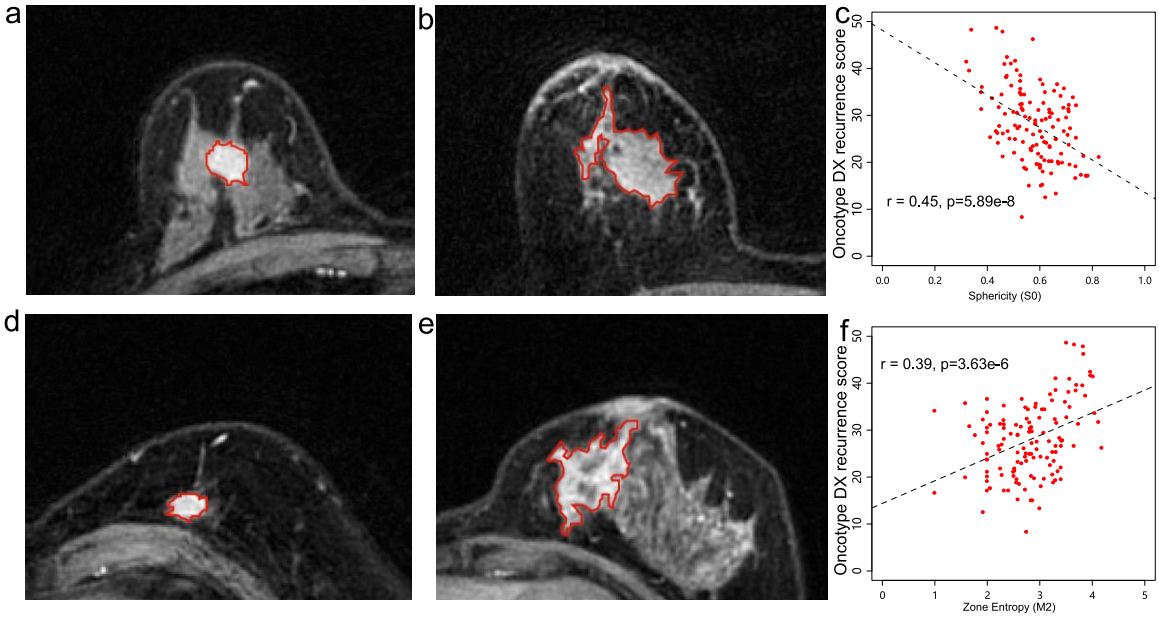

Figure 2. Examples of patients and distributions of radiogenomic signatures in the low- and high-risk groups of breast cancer patients. a) A 40-year-old patient with the following parameters: recurrence score (RS)=17.4; sphericity=0.784 and zone entropy=2.725. b) A 66-year-old patient with the following parameters: RS=38.5; sphericity=0.475 and zone entropy=3.691. c) The sphericity feature was negatively correlated with the RS. d) A 61-year-old patient with the following parameters: RS=17.2; sphericity=0.758 and zone entropy=2.322. e) A 32-year-old patient with the following parameters: RS= 41.5; sphericity=0.321 and zone entropy=4.018. f) The zone entropy feature was positively correlated with the RS. The illustrated images were obtained from the second postcontrast series.

Prognostic validation of the radiogenomic signatures

In this stage, the radiogenomic signatures established in the first stage were validated in the prediction of breast cancer patient survival using the prognostic validation cohort. The signatures that were identical to those in the radiogenomic cohort were used, with model parameters trained in the first stage to generate the pRS. The correlation map between the 11 radiogenomic signatures is illustrated in Supplementary Figure 3. The optimal threshold of the pRS for discriminating patients who achieved good survival from those who achieved poor survival was 29.9, which is close to the threshold of the clinically used Oncotype DX RS in categorizing high recurrence risk breast cancer (RS>30). Patients with a pRS less than 29.9 were associated with good OS (p=0.0014) or good RFS (p=0.0012) (Figure 3). These results indicate that the radiogenomic signatures derived from the genomic assay used to assess recurrence risk could be utilized to predict survival in breast cancer patients.

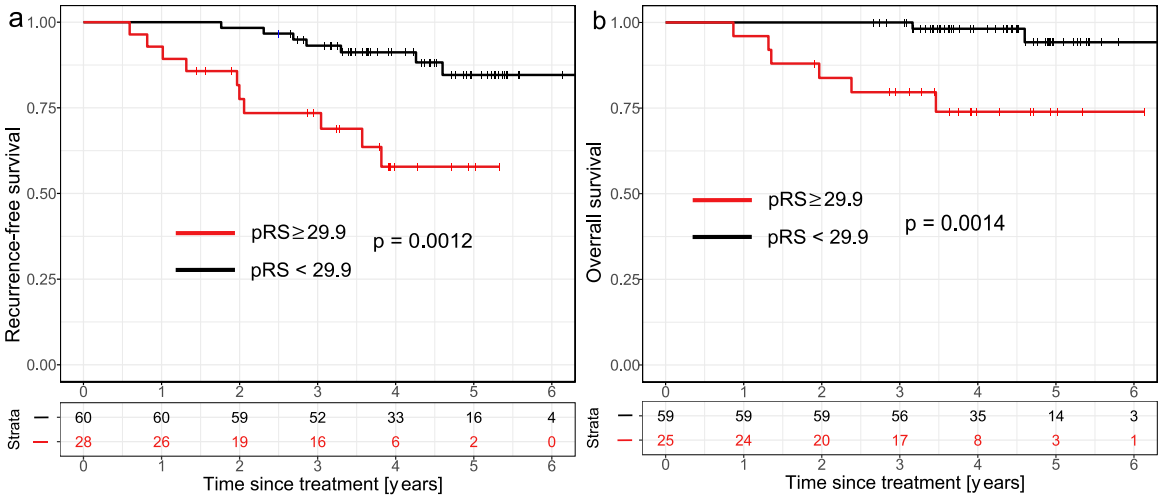

Figure 3. Kaplan–Meier survival analyses according to the pRS generated by the radiogenomic signatures for a) recurrence-free survival and b) overall survival.

Clinical information, including age, PR status and tumor size (longest diameter of the tumor), was added to control for these covariates in the predictive model. In the multivariate survival model, pRS remained independently correlated with either OS or RFS (p=0.021 and 0.002, respectively).

To enhance the predictive performance of the survival model using pRS alone, other imaging features that were correlated/complementary to the identified imaging signatures were incorporated in the model. Specifically, the imaging features that correlated with RFS/OS in the univariate survival analysis with p less than 0.1 were initially incorporated. After feature selection by using the step forward method, four additional features along with the pRS were combined in the multivariate model, which showed enhanced significance for the association with RFS (p=6.60e-05). Moreover, the multivariate survival model showed greater significance (7.64e-7) with OS upon the addition of five other imaging features. The pRS remained independently correlated with either OS or RFS (p=0.048 and 0.020, respectively) (Supplementary Tables 3 and 4) in the survival model. These results show that the prognostic radiogenomic signatures could be independently used and augment predictive power for the survival analysis of breast cancer patients.

Therapeutic validation of the radiogenomic signatures

In this stage, the therapeutic value of the radiogenomic signatures identified in the first stage was evaluated in the treatment validation cohort. The established regression model trained in the first stage was applied to this data cohort to generate pRS. The correlation map between the radiogenomic signatures is illustrated in Supplementary Figure 4. A significantly higher (p=0.001) pRS was observed in responders (n=38, mean pRS=30.51) than in nonresponders (n=97, mean pRS=27.35) (Figure 4a). A univariate SVM model using the pRS reached an AUC of 0.630. These results suggest that the radiognomic

signatures associated with recurrence biomarkers could be used as therapeutic biomarkers in breast cancer.

To enhance the prediction power, additional and complementary imaging features that showed low correlations with the pRS were incorporated in the model (Figure 4b). After the inclusion of these imaging features, the prediction performance was increased to an AUC of 0.6687. When pRS, the complementary imaging features and clinical information were combined, the model prediction performance was significantly (p=5.05E-05) increased to 0.849 (Figure 4b) compared to that using only pRS. These results suggest that the radiogenomic signatures derived from the prognostic genomic assay are promising in enhancing breast cancer treatment response prediction.

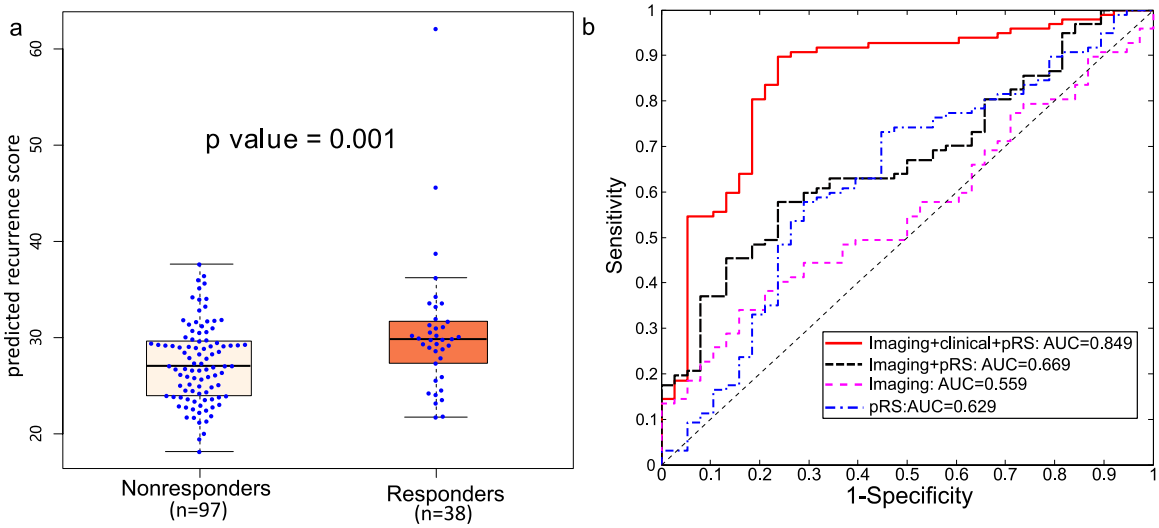

Figure 4. Therapeutic effectiveness of the pRS for discriminating responders from nonresponders. a) Boxplot of the pRS between the responder and nonresponder groups. b) ROC plots of the predictive models using pRS, a combination of pRS and other imaging features, a combination of pRS, clinical and other imaging features, and other imaging features.

Analysis of the radiogenomic signatures across the three cohorts

The radiogenomic signatures (n=11) were analyzed across the three cohorts for different targets (Figure 5). The texture features (e.g., difference average and zone entropy), which represent intratumor heterogeneity, were positively associated with the Oncotype DX RS. A high value of texture features was associated with poor RFS/OS and a poor response to NACT. On the other hand, the values of morphologic features (e.g., sphericity and flatness), which reflect roundness of tumor shape, were negatively correlated with the Oncotype DX RS (corrected p=0.023 and 0.071, respectively). Moreover, a high value of these features was associated with good patient survival and poor response to NACT. These results suggest consistent/robust trends of these imaging features correlated with genomic assays (Oncotype DX RS), responses to treatment, and survival.

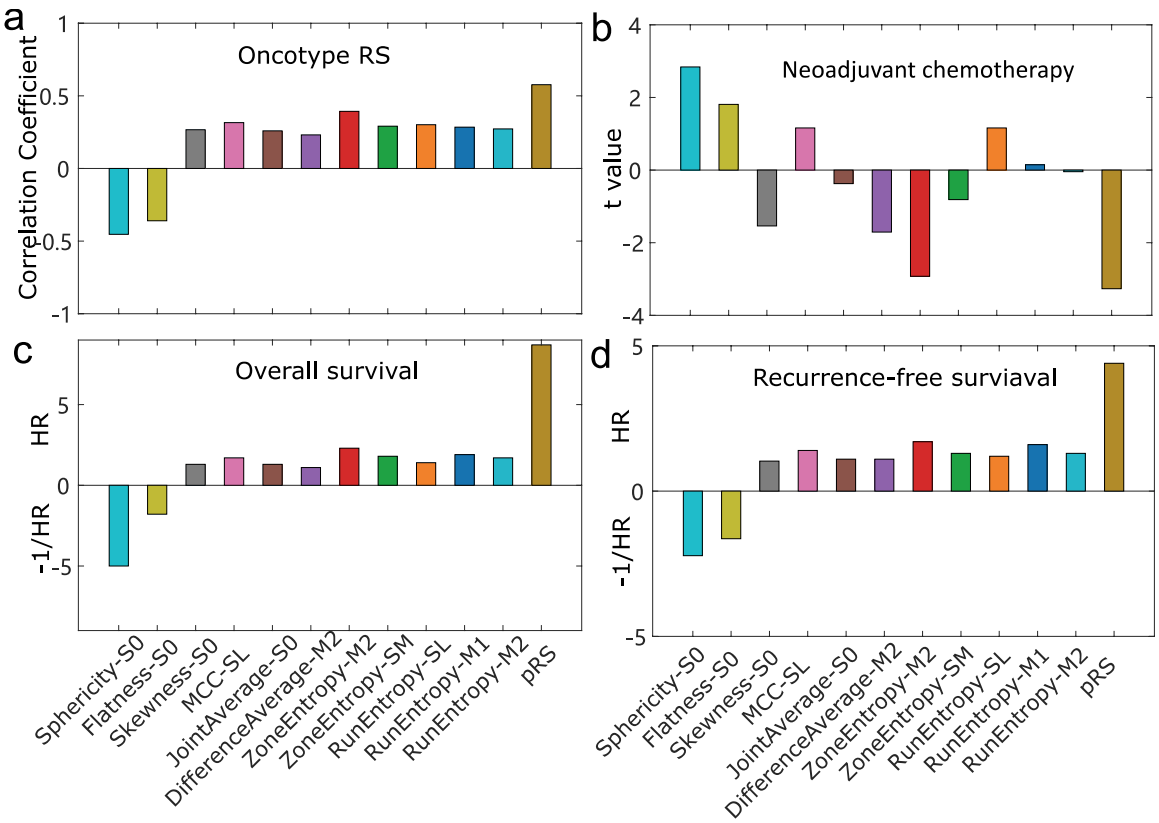

Figure 5. Bar plots show features that correlated with Oncotype DX RS, recurrence-free survival (RFS), overall survival (OS) and the response to NACT. For each feature, the Pearson correlation coefficient was used to measure the correlation with the Oncotype DX RS, the hazard ratio (HR) was used for survival (RFS and OS) analysis, and Student's t value was used to measure the difference in features between responders and nonresponders. In panels c and c, an HR value between 0 and 1 was transformed as -1/HR to better represent the trend of this measurement across different features.

Discussion

We proposed a strategy of radigenomic analysis by developing prognostic signatures associated with the Oncotype DX RS, which was subsequently validated in predicting the response to NACT and survival outcomes with multicohort analysis. The identified RS signatures were also used to enhance the prediction performance by combining complementary imaging features. The results demonstrated that the radiogenomic signatures derived from genomic assays could be used as independent and additional predictors for breast cancer diagnosis and treatment management. Our study provides a noninvasive biomarker that shows promising diagnostic and prognostic value in clinical practice.

Imaging provides an effective, noninvasive, and repeatable approach to improve treatment management and predict prognosis in breast cancer. Radiomic signatures, which reflect intrinsic imaging phenotypes of tumor heterogeneity at primary diagnosis, have been effectively used as predictors for recurrence [37] and the response to NACT in previous studies [38]. With large differences, we did not directly train a predictive model targeting survival or treatment responses in breast cancer patients but instead evaluated it

with two different while similar tasks. The results for pRS that can be validated across multiple cohorts may be explained by the fact that genomic assays (Oncotype DX RS) are regarded as prognostic and therapeutic biomarkers for breast cancer. More importantly, the predictive power of the pRS could be augmented by combining additional and complementary imaging features that correlated with the treatment responses and survival of breast cancer patients. Therefore, our study demonstrates that task-specific radiogenomic signatures can be used as noninvasive biomarkers that add more information to other related tasks for improved prognostic and therapeutic accuracy.

We included patients with node-positive breast cancer in the predictive model. Evolving evidence also suggests that the Oncotype RS has predictive value for node-positive breast cancer and that chemotherapy can be avoided in more than half of women with HR-positive breast cancer and 1 to 3 positive nodes [39]. Therefore, our machine learning model can be extended to patients with other statuses, such as those with node-positive HER2-positive breast cancer. The extendable model (also illustrated in the above paragraph) is, in principle, similar to transfer learning in the context of machine learning; that is, the knowledge from a trained source domain of recurrence biomarkers is transferred to enhance prediction in the target domain of different patient statuses or different clinical tasks (e.g., treatment and prognosis).

In our study, a high pRS was associated with a good response to NACT. A previous study showed that patients with a high Oncotype DX RS are more likely to benefit from chemotherapy than patients with a low RS [40]. As previously reported, patients with an intermediate Oncotype RS ($25 \leq RS \leq 30$) or a high RS ($RS > 30$) were associated with a larger absolute benefit of chemotherapy [41, 42]. These studies are in line with the findings of our study. The findings of our study illustrate that patients with a high pRS (greater than 29.9) had significantly poorer survival (OS and RFS) than the other patients. A previous study reported that patients with a low RS or an intermediate RS (18-30) had very low rates of distant recurrence and death, while the rates were high for those with a high RS [43], consistent with our results.

In this study, imaging features, including first-order, morphological and texture features, were analyzed. It was interesting to find that low or negative skewness was associated with a low RS. These results suggest that tumors with pixel value distributions of predominantly lower-than-average values tend to have a low risk of recurrence. A previous study showed a related finding [38], in which this negative value of skewness was associated with luminal A tumors; this tumor subtype is associated with a higher survival rate and a lower recurrence rate than other subtypes [44]. This result is partly consistent with our findings.

Our study showed that texture features, including the maximum correlation coefficient and difference average, were associated with survival in the multivariate model. The difference average feature measures the relationship between occurrences of pixel pairs with similar intensity values and occurrences of pairs with differing intensity values, which reflects the level of heterogeneity inside a tumor. The results suggested that a high

level of intratumor heterogeneity is associated with poor OS/RFS in breast cancer patients. An earlier study showed an association between the maximum correlation coefficient and the Oncotype DX RS, which is in line with our findings [45].

Additionally, morphologic features, including sphericity, which reflects tumor roundness, were mostly significant across the three cohorts. Tumor sphericity is associated with prognosis in oral cancer [46] and lung cancer [47]. A recent study showed that tumor sphericity could be used as an imaging predictor for pCR and the endpoint of the I-SPY 2 trial [48]. The authors reported higher tumor sphericity values in patients who achieved pCR than in patients who did not. These results are in line with our findings, i.e., tumors with an irregular shape (low sphericity) are likely associated with a high RS, no response to treatment and poor patient survival.

Limitations of our study should be addressed. First, the imaging data were obtained from three data cohorts with different imaging protocols, which may have induced variation in the imaging data and eventually the radiomic analysis of tumor images. To reduce bias in the imaging data, we performed careful quality control by excluding low-quality images. Proper image preprocessing and normalization were also conducted to control for bias among the imaging protocols. Second, radiomic analysis of the parenchyma region should be performed to facilitate a comprehensive analysis of imaging heterogeneity in breast cancer. Third, multiparametric images including the apparent diffusion coefficient derived from DWI should also be examined, which may help to identify patients with the highest recurrence risk and those who would benefit most from chemotherapy [49, 50].

In conclusion, our results demonstrated that radiogenomic signatures can be used as independent and additional value in the prognosis and treatment of breast cancer by multicohort analysis. This strategy of radiogenomic analysis identifies noninvasive features of tumor heterogeneity that are consistently correlated with multiple targets in breast cancer and could serve as a promising, noninvasive biomarker in future clinical practice.

References

[1] P. Rastogi, S. J. Anderson, H. D. Bear, C. E. Geyer, M. S. Kahlenberg, A. Robidoux *et al.*, Preoperative chemotherapy: updates of National Surgical Adjuvant Breast and Bowel Project Protocols B-18 and B-27, *J Clin Oncol* **2008**; 26: 778-85.

[2] M. Golshan, C. T. Cirrincione, W. M. Sikov, D. A. Berry, S. Jasinski, T. F. Weisberg *et al.*, Impact of neoadjuvant chemotherapy in stage II-III triple negative breast cancer on eligibility for breast-conserving surgery and breast conservation rates: surgical results from CALGB 40603 (Alliance), *Ann Surg* **2015**; 262: 434-9; discussion 438-9.

[3] L. M. Spring, G. Fell, A. Arfe, C. Sharma, R. Greenup, K. L. Reynolds *et al.*, Pathologic Complete Response after Neoadjuvant Chemotherapy and Impact on Breast Cancer Recurrence and Survival: A Comprehensive Meta-analysis, *Clin Cancer Res* **2020**; 26: 2838-2848.

[4] R. Chen, Y. Ye, C. Yang, Y. Peng, B. Zong, F. Qu *et al.*, Assessment of the predictive role of pretreatment Ki-67 and Ki-67 changes in breast cancer patients receiving neoadjuvant chemotherapy according to the molecular classification: a retrospective study of 1010 patients, *Breast Cancer Research and Treatment* **2018**; 170: 35-43.

[5] J. S. Parker, M. Mullins, M. C. Cheang, S. Leung, D. Voduc, T. Vickery *et al.*, Supervised risk predictor of breast cancer based on intrinsic subtypes, *J Clin Oncol* **2009**; 27: 1160-7.

[6] M. J. van de Vijver, Y. D. He, L. J. van't Veer, H. Dai, A. A. Hart, D. W. Voskuil *et al.*, A gene-expression signature as a predictor of survival in breast cancer, *N Engl J Med* **2002**; 347: 1999-2009.

[7] L. J. van 't Veer, H. Dai, M. J. van de Vijver, Y. D. He, A. A. Hart, M. Mao *et al.*, Gene expression profiling predicts clinical outcome of breast cancer, *Nature* **2002**; 415: 530-6.

[8] S. Paik, S. Shak, G. Tang, C. Kim, J. Baker, M. Cronin *et al.*, A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer, *N Engl J Med* **2004**; 351: 2817-26.

[9] V. I. Petkov, D. P. Miller, N. Howlader, N. Gliner, W. Howe, N. Schussler *et al.*, Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study, *NPJ Breast Cancer* **2016**; 2: 16017.

[10] A. M. Pease, L. A. Riba, R. A. Gruner, N. M. Tung, and T. A. James, Oncotype DX® Recurrence Score as a Predictor of Response to Neoadjuvant Chemotherapy, *Annals of Surgical Oncology* **2019**; 26: 366-371.

[11] S. Harnan, P. Tappenden, K. Cooper, J. Stevens, A. Bessey, R. Rafia *et al.*, Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer: a systematic review and economic analysis, *Health Technol Assess* **2019**; 23: 1-328.

[12] W. Wang, L. Lin, X. Fei, J. Hong, W. Gao, S. Zhu *et al.*, 21-Gene recurrence score influences the chemotherapy decision for patients with breast cancer of different luminal subtypes, *Oncol Lett* **2019**; 18: 4346-4356.

[13] S. G. Wu, W. W. Zhang, J. Wang, Y. Dong, Y. X. Chen, and Z. Y. He, Effect of 21-gene recurrence score in decision-making for surgery in early stage breast cancer, *Onco Targets Ther* **2019**; 12: 2071-2078.

[14] M. A. Rabie, A. Rankin, A. Burger, and M. Youssef, The effect of Oncotype DX((R)) on adjuvant chemotherapy treatment decisions in early breast cancer, *Ann R Coll Surg Engl* **2019**; 101: 596-601.

[15] A. Olsson-Brown, P. Piskilidis, J. O'Hagan, N. Thorp, P. Robson, H. Innes *et al.*, The impact of the 21-gene recurrence score (Oncotype DX) on concordance of adjuvant therapy decision making as measured by the Liverpool Systemic Therapy Adjuvant Decision Tool, *Breast* **2019**; 44: 94-100.

[16] L. Zhang, M. C. Hsieh, V. Petkov, Q. Yu, Y. W. Chiu, and X. C. Wu, Trend and survival benefit of Oncotype DX use among female hormone receptor-positive breast cancer patients in 17 SEER registries, 2004-2015, *Breast Cancer Res Treat* **2020**; 180: 491-501.

[17] S. Park, Y. Han, Y. Liu, A. T. Toriola, L. L. Peterson, G. A. Colditz *et al.*, Adjuvant chemotherapy and survival among patients 70 years of age and younger with node-negative breast cancer and the 21-gene recurrence score of 26-30, *Breast Cancer Res* **2019**; 21: 110.

[18] P. P. Peethambaram, T. L. Hoskin, C. N. Day, M. P. Goetz, E. B. Habermann, and J. C. Boughey, Use of 21-gene recurrence score assay to individualize adjuvant chemotherapy recommendations in ER+/HER2- node positive breast cancer-A National Cancer Database study, *NPJ Breast Cancer* **2017**; 3: 41.

[19] C. B. Weldon, J. R. Trosman, W. J. Gradishar, A. B. Benson, 3rd, and J. C. Schink, Barriers to the use of personalized medicine in breast cancer, *J Oncol Pract* **2012**; 8: e24-31.

[20] D. T. Tsoi, M. Inoue, C. M. Kelly, S. Verma, and K. I. Pritchard, Cost-effectiveness analysis of recurrence score-guided treatment using a 21-gene assay in early breast cancer, *Oncologist* **2010**; 15: 457-65.

[21] S. H. Yoo, T. Y. Kim, M. Kim, K. H. Lee, E. Lee, H. B. Lee *et al.*, Development of a Nomogram to Predict the Recurrence Score of 21-Gene Prediction Assay in Hormone Receptor-Positive Early Breast Cancer, *Clin Breast Cancer* **2020**; 20: 98-107 e1.

[22] K. J. Nam, H. Park, E. S. Ko, Y. Lim, H. H. Cho, and J. E. Lee, Radiomics signature on 3T dynamic contrast-enhanced magnetic resonance imaging for estrogen receptor-positive invasive breast cancers: Preliminary results for correlation with Oncotype DX recurrence scores, *Medicine (Baltimore)* **2019**; 98: e15871.

[23] A. Saha, M. R. Harowicz, W. Wang, and M. A. Mazurowski, A study of association of Oncotype DX recurrence score with DCE-MRI characteristics using multivariate machine learning models, *J Cancer Res Clin Oncol* **2018**; 144: 799-807.

[24] G. A. Woodard, K. M. Ray, B. N. Joe, and E. R. Price, Qualitative Radiogenomics: Association between Oncotype DX Test Recurrence Score and BI-RADS Mammographic and Breast MR Imaging Features, *Radiology* **2018**; 286: 60-70.

[25] J. Y. Kim, J. J. Kim, L. Hwangbo, J. W. Lee, N. K. Lee, K. J. Nam *et al.*, Diffusion-weighted MRI of estrogen receptor-positive, HER2-negative, node-negative breast cancer: association between intratumoral heterogeneity and recurrence risk, *Eur Radiol* **2020**; 30: 66-76.

[26] M. Fan, P. P. Xia, R. Clarke, Y. Wang, and L. H. Li, Radiogenomic signatures reveal multiscale intratumour heterogeneity associated with biological functions and survival in breast cancer, *Nature Communications* **2020**; 11: 12.

[27] R. D. Chitalia, and D. Kontos, Role of texture analysis in breast MRI as a cancer biomarker: A review, *J Magn Reson Imaging* **2019**; 49: 927-938.

[28] S. Paik, G. Tang, S. Shak, C. Kim, J. Baker, W. Kim *et al.*, Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer, *J Clin Oncol* **2006**; 24: 3726-34.

[29] K. N. Ogston, I. D. Miller, S. Payne, A. W. Hutcheon, T. K. Sarkar, I. Smith *et al.*, A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival, *The Breast* **2003**; 12: 320-327.

[30] Q. Zhu, L. Q. Wang, S. Tannenbaum, A. Ricci, P. DeFusco, and P. Hegde, Pathologic response prediction to neoadjuvant chemotherapy utilizing pretreatment near-infrared imaging parameters and tumor pathologic criteria, *Breast Cancer Research* **2014**; 16:

[31] M. E. Hammond, D. F. Hayes, A. C. Wolff, P. B. Mangu, and S. Temin, American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer, *J Oncol Pract* **2010**; 6: 195-7.

[32] A. C. Wolff, M. E. Hammond, D. G. Hicks, M. Dowsett, L. M. McShane, K. H. Allison *et al.*, Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update, *J Clin Oncol* **2013**; 31: 3997-4013.

[33] M. Fan, Z. H. Liu, S. D. Xie, M. S. Xu, S. W. Wang, X. Gao *et al.*, Integration of dynamic contrast-enhanced magnetic resonance imaging and T2-weighted imaging radiomic features by a canonical correlation analysis-based feature fusion method to predict histological grade in ductal breast carcinoma, *Physics in Medicine and Biology* **2019**; 64: 13.

[34] Q. Yang, L. Li, J. Zhang, G. Shao, C. Zhang, and B. Zheng, Computer-aided diagnosis of breast DCE-MRI images using bilateral asymmetry of contrast enhancement between two breasts, *J Digit Imaging* **2014**; 27: 152-60.

[35] J. J. M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan *et al.*, Computational Radiomics System to Decode the Radiographic Phenotype, *Cancer Res* **2017**; 77: e104-e107.

[36] H. Zou, and T. Hastie, Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **2005**; 67: 301-320.

[37] R. D. Chitalia, J. Rowland, E. S. McDonald, L. Pantalone, E. A. Cohen, A. Gastounioti *et al.*, Imaging Phenotypes of Breast Cancer Heterogeneity in Preoperative Breast Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) Scans Predict 10-Year Recurrence, *Clin Cancer Res* **2020**; 26: 862-869.

[38] M. Fan, G. Wu, H. Cheng, J. Zhang, G. Shao, and L. Li, Radiomic analysis of DCE-MRI for prediction of response to neoadjuvant chemotherapy in breast cancer patients, *Eur J Radiol* **2017**; 94: 140-147.

[39] S. Ahmed, S. Pati, D. Le, K. Haider, and N. Iqbal, The prognostic and predictive role of 21-gene recurrence scores in hormone receptor-positive early-stage breast cancer, *J Surg Oncol* **2020**;

[40] N. Green, A. Al-Allak, and C. Fowler, Benefits of introduction of Oncotype DX((R)) testing, *Ann R Coll Surg Engl* **2019**; 101: 55-59.

[41] A. F. Ibraheem, D. J. Press, O. I. Olopade, and D. Huo, Community clinical practice patterns and mortality in patients with intermediate oncotype DX recurrence scores: Who benefits from chemotherapy?, *Cancer* **2019**; 125: 213-222.

[42] C. E. Geyer, Jr., G. Tang, E. P. Mamounas, P. Rastogi, S. Paik, S. Shak *et al.*, 21-Gene assay as predictor of chemotherapy benefit in HER2-negative breast cancer, *NPJ Breast Cancer* **2018**; 4: 37.

[43] S. M. Stemmer, M. Steiner, S. Rizel, L. Soussan-Gutman, N. Ben-Baruch, A. Bareket-Samish *et al.*, Clinical outcomes in patients with node-negative breast cancer treated based on the recurrence score results: evidence from a large prospectively designed registry, *NPJ Breast Cancer* **2017**; 3: 33.

[44] K. D. Voduc, M. C. U. Cheang, S. Tyldesley, K. Gelmon, T. O. Nielsen, and H. Kennecke, Breast Cancer Subtypes and the Risk of Local and Regional Relapse, *Journal of Clinical Oncology* **2010**; 28: 1684-1691.

[45] H. Li, Y. Zhu, E. S. Burnside, K. Drukker, K. A. Hoadley, C. Fan *et al.*, MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays, *Radiology* **2016**; 281: 382-391.

[46] A. Tarsitano, F. Ricotta, L. Cercenelli, B. Bortolani, S. Battaglia, E. Lucchi *et al.*, Pretreatment tumor volume and tumor sphericity as prognostic factors in patients with oral cavity squamous cell carcinoma, *J Craniomaxillofac Surg* **2019**; 47: 510-515.

[47] A. Davey, M. van Herk, C. Faivre-Finn, H. Mistry, and A. McWilliam, Is tumour sphericity an important prognostic factor in patients with lung cancer?, *Radiother Oncol* **2020**; 143: 73-80.

[48] W. Li, D. C. Newitt, B. Yun, E. F. Jones, V. Arasu, L. J. Wilmes *et al.*, Tumor Sphericity Predicts Response in Neoadjuvant Chemotherapy for Invasive Breast Cancer, *Tomography* **2020**; 6: 216-222.

[49] N. Amornsiripanitch, V. T. Nguyen, H. Rahbar, D. S. Hippe, V. K. Gadi, M. H. Rendi *et al.*, Diffusion-weighted MRI characteristics associated with prognostic pathological factors and recurrence risk in invasive ER+/HER2- breast cancers, *J Magn Reson Imaging* **2018**; 48: 226-236.

[50] S. B. Thakur, M. Durando, S. Milans, G. Y. Cho, L. Gennaro, E. J. Sutton *et al.*, Apparent diffusion coefficient in estrogen receptor-positive and lymph node-negative invasive breast cancers at 3.0T DW-MRI: A potential predictor for an oncotype Dx test recurrence score, *J Magn Reson Imaging* **2018**; 47: 401-409.