Medicine®

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

Kyung Jin Nam, MD[a], Hyunjin Park, PhD[b,c], Eun Sook Ko, MD, PhD[d,*], Yaeji Lim, PhD[e], Hwan-Ho Cho, BS[f], Jeong Eon Lee, MD, PhD[g]

Abstract

To evaluate the ability of a radiomics signature based on 3T dynamic contrast-enhanced (DCE) magnetic resonance imaging (MRI) to distinguish between low and non-low Oncotype DX (OD) risk groups in estrogen receptor (ER)-positive invasive breast cancers.

Between May 2011 and March 2016, 67 women with ER-positive invasive breast cancer who performed preoperative 3T MRI and OD assay were included. We divided the patients into low (OD recurrence score [RS] <18) and non-low risk (RS ≥18) groups. Extracted radiomics features included 8 morphological, 76 histogram-based, and 72 higher-order texture features. A radiomics signature (Rad-score) was generated using the least absolute shrinkage and selection operator (LASSO). Univariate and multivariate logistic regression analyses were performed to investigate the association between clinicopathologic factors, MRI findings, and the Rad-score with OD risk groups, and the areas under the receiver operating characteristic curves (AUC) were used to assess classification performance of the Rad-score.

The Rad-score was constructed for each tumor by extracting 10 (6.3%) from 158 radiomics features. A higher Rad-score (odds ratio [OR], 65.209; $P<.001$), Ki-67 expression (OR, 17.462; $P=.007$), and high p53 (OR=8.449; $P=.077$) were associated with non-low OD risk. The Rad-score classified low and non-low OD risk with an AUC of 0.759.

The Rad-score showed the potential for discrimination between low and non-low OD risk groups in patients with ER-positive invasive breast cancers.

Abbreviations: AUC = area under the curve, BI-RADS = Breast Imaging Reporting and Data System, CAD = computer-aided diagnosis, DCE = dynamic contrast-enhanced, EIC = extensive intraductal component, ER = estrogen receptor, LASSO = logistic least absolute shrinkage and selection operator, LVI = lymphovascular invasion, MRI = magnetic resonance imaging, OD = Oncotype DX, PR = progesterone receptor, ROC = receiver operating characteristic, ROI = region of interest, RS = recurrence Score.

Keywords: breast, breast cancer, magnetic resonance imaging, Oncotype DX, radiomics

1. Introduction

The Oncotype DX (OD) genomic assay (Oncotype DX, Genomic Health, CA) evaluates the expression of a panel of 21 genes from a tumor specimen. It is a validated multigene diagnostic assay that predicts the likelihood of breast cancer recurrence and whether patients are likely to benefit from commonly-used chemotherapy regimens in estrogen receptor (ER)-positive early breast cancer patients.[1–9] The assay results are combined into a single score, or

recurrence score (RS). According to Paik et al,[2] the rate of distant recurrence at 10 years in patients with a low RS (6.8%) was significantly lower than in those with a high RS (30.5%), while another study showed that RS predicts the magnitude of chemotherapy benefit.[3] Thus, in their treatment guidelines for early breast cancer, the American Society of Clinical Oncology strongly recommends chemotherapy in patients with a high RS (>30), not low RS (<18).[10] However, the OD assay is performed on tissue samples, involves additional tumor handling, and is expensive.[11,12] Therefore, a non-invasive means of distinction in patients who may be treated without chemotherapy would facilitate treatment decision-making, cost-reduction, and chemotherapy-induced toxicity avoidance.

Radiomics is a field that converts imaging data into high-dimensional, mineable, and quantitative features, including characteristics that are imperceptible to the human eye.[13–15] Radiomics research with genomic information, or "radiogenomics", is a relatively new field of research.[16] The aim of radiogenomics is to correlate imaging characteristics (i.e., the imaging phenotype) with underlying gene expression patterns, gene mutations, and other genome-related characteristics.[16] Radiogenomics literature on brain, breast, and lung cancers, among others, is growing.[16–20] In radiogenomic research in breast imaging, correlation with OD might be representative topic. There are several studies that have identified radiomics features on magnetic resonance imaging (MRI) that might be associated with OD RS. Sutton et al[15] reported that morphological and texture-based MRI features correlated with OD RS, while another study showed that 31 extracted multi-parametric MRI (1.5T) features correlated with OD RS.[21] However, both studies used a limited number of features to derive a predictive biomarker. While the features, as individual predictors, were significantly associated with prognosis prediction in patients with breast cancer, the combined analysis of a panel of multiple predictors as a signature is more powerful for assisting clinical management.[22,23] Radiomics enables multiple imaging features to be investigated in parallel and thus facilitates the construction of powerful signatures.[13,14] Recently, radiomics signatures were found to be predictive in non-small cell lung cancer and colorectal cancer.[24–26] To our knowledge, the correlation between radiomics signature and OD risk classification is currently unclear.

Therefore, the purpose of this study was to evaluate the ability of a radiomics signature based on 3T dynamic contrast-enhanced (DCE)-MRI to distinguish between low and non-low OD risk groups in patients with ER-positive invasive breast cancers.

2. Materials and methods

2.1. Study population

Our Institutional Review Board approved this retrospective study and informed consent was waived.

We identified 127 women with ER-positive invasive breast cancer who underwent preoperative breast MRI and a validated gene expression assay (Oncotype DX; Genomic Health, Redwood City, Calif) at our institution, from May 2011 to March 2016. The study inclusion criteria were:

(1) completion of preoperative 3T DCE-MRI at our institution,
(2) initial unilateral breast malignancy, and
(3) a lesion manifesting as a mass at MRI. Patients who already had a malignancy in another organ (metastasis or primary malignancy) were excluded. Finally, the study population

included 67 women (mean age, 45.1 years; range, 28–66 years).

2.2. Magnetic resonance imaging protocol

MRI was performed using a 3T Achieva scanner (Philips Medical Systems, Best, The Netherlands), with a dedicated bilateral phase array breast coil, in the prone position. The MRI examination consisted of turbo spin-echo T1- and T2- weighted sequences, and a 3-dimensional DCE sequence. Details of sequence parameters are described in Appendix S1, http://links.lww.com/MD/D22. For dynamic contrast enhancement, a 0.1 mmol/kg bolus of Gadobutrol (Gadovist; Bayer Healthcare Pharmaceutical, Berlin, Germany) was injected, followed by a 20-mL saline flush.

2.3. Magnetic resonance imaging preparation for radiomics analysis

T2-weighted, pre-enhanced T1-weighted, contrast-enhanced T1-weighted and contrast-enhanced T1-weighted subtraction MR images were retrieved from the Picture Archiving Communication System and loaded onto a workstation for further radiomics analysis. Subtraction images, from contrast-enhanced images at 90 s to pre-enhanced images and contrast-enhanced T1-weighted images at 90 s after contrast injection, were assessed. The images were reviewed by a radiologist with 12 years of experience in breast MRI (ESK.). In every slice, the radiologist drew a region of interest (ROI) around the visible tumor on the contrast-enhanced T1-weighted subtraction images. Finally, the ROIs covered the entire 3D volume of the tumor for volumetric segmentation. The defined ROI was co-registered onto the other 3 image series using a 9-parameter affine transform with mutual information as the similarity measure.[27] The ROI was drawn as large as possible but did not include edge voxels to avoid a partial volume effect. In multifocal or multicentric disease, the largest tumor was selected as an index cancer for analysis.

2.4. Conventional magnetic resonance imaging and clinicopathological evaluation

The MR images of included masses were retrospectively evaluated, according to the American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) MR lexicon,[28] by 2 board-certified radiologists (JSC and ESK, with 8 and 12 years of experience in breast MR imaging, respectively) in consensus. The radiologists assessed the shape (oval/round, or irregular), margin (circumscribed, irregular, or spiculated), and internal enhancement characteristics (homogeneous, heterogeneous, or rim) of each mass.

All T1-weighted images were transferred to a computer-aided diagnosis (CAD) system (CAD stream, version 4.1.3, Merge Healthcare, Chicago, IL), a commercially available CAD system. Pixel values at the pre-contrast and second contrast-enhanced series (90 s after contrast injection) were compared to classify enhancement. If the pixel value increased above a threshold of 50%, it was shown in color. Color was assigned according to changes in pixel values, between the second contrast-enhanced and delayed contrast-enhanced series, as follows: persistent type indicated an increased pixel signal intensity greater than 10% from the second post-contrast series; washout type indicated decreased pixel signal intensity at the last post-contrast series

greater than 10% from the second post-contrast series; plateau type indicated change in neither direction by more than 10%. We chose the most suspicious kinetic curve component within the tumors. CAD report was prospectively saved by a research assistant. The following parameters were used in our analysis: peak enhancement value (%), most suspicious enhancement kinetics of delayed phase (persistent, plateau, or washout), and percentage of washout.

The final histopathologic results of surgical specimens were reviewed to determine tumor type, size, histologic grade, presence of an extensive intraductal component (EIC), lymphovascular invasion (LVI), ER, progesterone receptor (PR), Ki-67, and p53 status. ER and PR positivity were determined by using a cutoff value of less than 1% positively stained nuclei. We reviewed the T stage and N stage after surgery.

2.5. Statistical analysis

Statistical analyses were performed using R software (version 3.2.4; R Foundation for Statistical Computing, Vienna, Austria), and statistical significance was accepted where P values were < .05. The minimum required sample size was 14 which was computed with 2 tailed significance of 0.05 and power of 0.95 to detect RS score change of 10.[29]

2.6. Patient characteristics according to Oncotype DX recurrence score risk group

Patient characteristics, including clinicopathologic (age at diagnosis, tumor type, tumor size, histologic grade, presence of EIC, LVI, ER, PR, Ki-67, p53, T stage, and N stage), MRI findings according to BI-RADS lexicon (tumor shape, margin, and internal enhancement characteristics), and kinetic features assessed by CAD (the most suspicious enhancement kinetics of delayed phase, peak enhancement value [%], and percentage of washout) among low (RS < 18), intermediate (RS 18–30), and high (RS > 30) OD risk groups were compared using analysis of variance (ANOVA) for continuous variables, and Chi-square and Fisher exact tests for categorical variables.

2.7. Radiomics analysis and score building

Radiomics features from 4 MRI series were automatically computed over radiologist-drawn ROIs. The features were computed using a combination of open source code[30] and in-house generated computer code implemented in MATLAB (Mathworks Inc. Natick, MA). For most features, we used the open source software PyRadiomics so that the results could be easily reproduced. For the features not available in PyRadiomics, we used our in-house MATLAB code. The in-house code is attached as the supplement. The features were grouped into morphological (8 features), histogram-based (19 features), and higher-order texture features (18 features). A total of 158 features in 3 different categories were computed. In Appendix S2, http://links.lww.com/MD/D22 and Appendix Table 1, http://links.lww.com/MD/D22, categorical concepts of radiomics features (Appendix S2, http://links.lww.com/MD/D22) and the mathematical definition of adopted feature algorithms (Appendix Table 1, http://links.lww.com/MD/D22) are described. The features were z-score normalized. We adopted the least absolute shrinkage and selection operator (LASSO) method, a popular regularized machine learning method for high-dimensional data,

to identify significant radiomics features related to the primary outcome. The LASSO encourages the selected features to be orthogonal. A leave-1-out cross validation (LOOCV) was adopted to separate train and test data. Each iteration of LOOCV led to possibly different models and we kept the features and models that appeared more than 55% during the LASSO procedure to identify the stable features and models. We applied the user chosen threshold of 55% obtain a few stable features. Lowering the threshold might lead to too many unstable features and increasing the threshold might lead to having no common features. Our choice was a middle ground between 2 extremes. A radiomics signature (Rad-score) was calculated for each patient via a linear combination of selected radiomics features weighted by their respective coefficients. The coefficients were the mean values of the stable models.

2.8. Validated factors selection for predicting Oncotype DX recurrence score

Univariate logistic regression analyses were conducted with clinicopathologic, MRI findings according to BI-RADS lexicon, kinetic features assessed on CAD, and Rad-score, to find the significant factors for classifying low and non-low (intermediate and high) OD risk groups. The variables with P values < .05 were included in the multivariate logistic regression model. To avoid multicollinearity, we further performed stepwise selection using Akaike information criterion (AIC).

Receiver operating characteristic (ROC) curves using LOOCV were used to compare the performances of the Rad-score only, clinicopathologic, and combined models for predicting low and non-low OD risk groups. We computed area under the curve (AUC) values for numerical comparison.

The overall scheme of our pipeline is given in Figure 1.

3. Results

3.1. Patient characteristics according to the Oncotype DX recurrence score risk groups

Of the 67 patients, 45 (67.2%) were classified as low-risk, 19 (28.4%) intermediate-risk, and 3 (4.5%) high-risk. We further classified into 2 groups: low-risk (45/67, 67.2%), and non-low-risk (22/67, 32.8%). The median OD RS was 14 (range 3–39).

Patients' characteristics according to OD risk are summarized in Table 1. All 67 patients were ER-positive and HER2-negative. Among the clinical, pathologic, and radiologic parameters, a higher histologic grade (P < .001), PR negativity (P < .001), high Ki-67 expression (P < .001), p53 overexpression (P < .001), and oval/round shape (P = .038) were associated with higher OD RS. In 59 (88.1%) of the 67 MRI examinations, CAD reports were available.

3.2. Radiomics analysis and Rad-score building

One hundred fifty-six radiomics features were reduced to 10 (6.2%) potential predictors. The radiomics features with a nonzero coefficient in the LASSO regression model were as follows: 6 histogram-based features (Min_val_T2, Skewness_PRE T1, Skewness_Sub T1, Skewness_T2, Energy_val_Sub T1, and percentile_histogram.2.5_T2) and 4 higher-order texture features (Contrast_glcm_val_T2, Max_probability_val_Sub T1, Variance_glcm_val_CE T1, and Sum_variance_val_PRE T1).
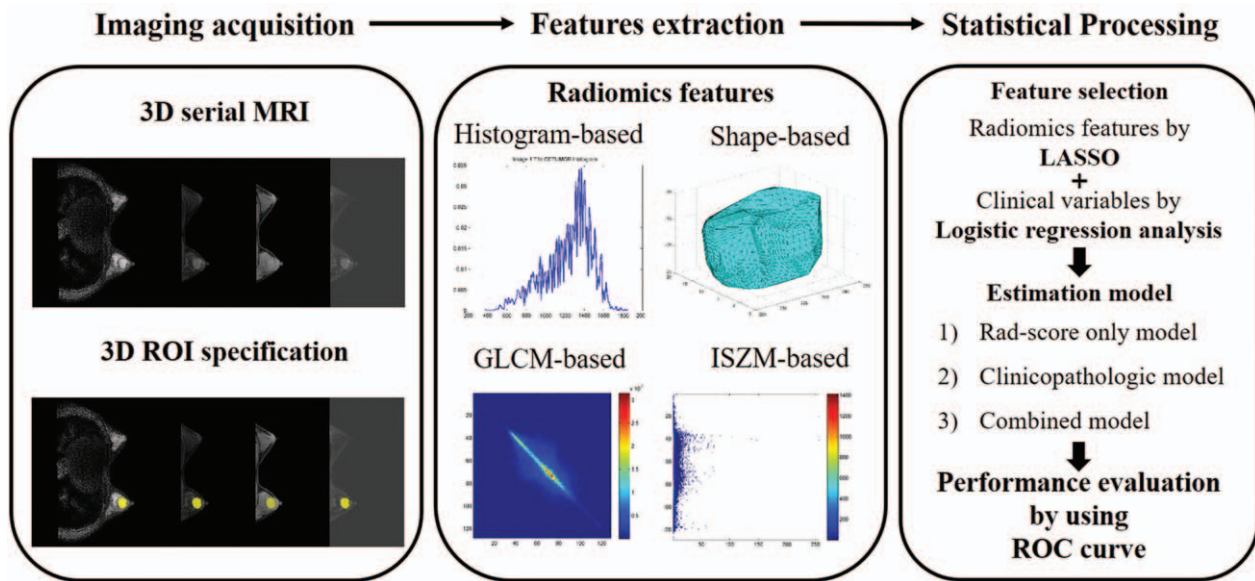
Figure 1. The overall scheme of the pipeline for present study.

Table 1

Characteristics of 67 patients with estrogen receptor (ER)-positive invasive breast cancers according to Oncotype DX risk groups.

	Oncotype DX recurrence score			
Variable	Low (n=45)	Intermediate (n=19)	High (n=3)	P value
Recurrence score	12.00±3.26	22.74±3.87	35.67±2.89	<.001
Patient age, yr*	45.29±8.00	44.00±6.82	49.33±10.26	.526
Tumor type				.721
IDC	41 (91.1)	19 (100.0)	3 (100.0)	
ILC	3 (6.7)	0 (0.0)	0 (0.00)	
Mucinous carcinoma	1 (2.2)	0 (0.0)	0 (0.00)	
Tumor size, cm*	1.70±0.69	1.73±0.86	1.10±0.87	.386
Histologic grade				<.001
1	24 (53.3)	4 (21.1)	0 (0.00)	
2	20 (44.4)	14 (73.7)	0 (0.00)	
3	1 (2.2)	1 (5.3)	3 (100.0)	
Extensive intraductal component				.234
Absent	29 (64.4)	15 (78.9)	1 (33.3)	
Present	16 (35.6)	4 (21.1)	2 (66.7)	
Lymphovascular invasion				.438
Absent	29 (64.4)	12 (63.2)	3 (100.0)	
Present	16 (35.6)	7 (36.8)	0 (0.0)	
PR				<.001
Negative	0 (0.0)	2 (10.5)	2 (66.7)	
Positive	45 (100.0)	17 (89.5)	1 (33.3)	
Ki-67				<.001
<14%	37 (82.2)	8 (42.1)	0 (0.0)	
≥14%	8 (17.8)	11 (57.9)	3 (100.0)	
p53				<.001
Negative	43 (95.6)	14 (73.7)	0 (0.0)	
Positive	2 (4.4)	5 (26.3)	3 (100.0)	
T stage				.870
1	35 (77.8)	14 (73.7)	2 (66.7)	
2	10 (22.2)	5 (26.3)	1 (33.3)	
N stage				.596
0	35 (77.8)	14 (73.7)	3 (100.0)	
1	10 (22.2)	5 (26.3)	0 (0.0)	

(continued)

Table 1

(continued).

Variable	Oncotype DX recurrence score			P value
	Low (n = 45)	Intermediate (n = 19)	High (n = 3)	
Tumor shape				.038
Oval or round	19 (42.2)	13 (68.4)	0 (0.0)	
Irregular	26 (57.8)	6 (31.6)	3 (100.0)	
Tumor margin				.334
Circumscribed	20 (44.4)	12 (63.2)	1 (33.3)	
Not circumscribed	25 (55.6)	7 (36.8)	2 (66.7)	
Internal enhancement				.054
Homogeneous	32 (71.1)	12 (63.2)	2 (66.7)	
Heterogeneous	10 (22.2)	1 (5.3)	1 (33.3)	
Rim enhancement	3 (6.7)	6 (31.6)	0 (0.0)	
Enhancement kinetics of delayed phase (n = 59)				.616
Persistent	1 (2.5)	0 (0.0)	0 (0.0)	
Plateau	1 (2.5)	2 (11.8)	0 (0.0)	
Washout	38 (95.0)	15 (88.2)	2 (100.0)	
Peak enhancement value[*]	347.15 ± 160.81	322.71 ± 124.64	471.00 ± 111.72	.416
Percentage of washout[*]	31.20 ± 25.19	28.12 ± 23.07	18.50 ± 17.68	.729

IDC = invasive ductal carcinoma, ILC = invasive lobular carcinoma, PR = progesterone receptor, RS = recurrence score.

[*] Data are means ± standard deviation, Data in parentheses are percentages.

The radiomics signature was constructed with a Rad-score, calculated using the following formula:

Rad-score = Min_val_T2 x 0.0259 + Skewness_T2 x 0.0903 + percentile_histogram.2.5_T2 x 0.0798 + Contrast_glcm_val_T2 x (−0.0048) + Skewness_Sub T1 x 0.0055 + Energy_val_Sub T1 x (−0.0708) + Max_probability_val_Sub T1 x 0.0495 + Skewness_PRE T1 x (−0.0608) + Sum_variance_val_PRE T1 x 0.0399 + Variance_glcm_val_CE T1 x (−0.0190).

The median Rad-score yielded -0.0676 (range, −0.2439–0.1758; interquartile range, −0.1564–0.0352) in 45 low-risk patients and 0.1197 (range, −0.1681–0.4245; interquartile range, 0.0013–0.2521) in 22 non-low risk patients. The median Rad-score of 67 patients was −0.0258 (range, −0.2439–0.4245; interquartile range, −0.1262–0.0985).

3.3. Prediction models for Oncotype DX recurrence score: buildings and performances

In univariate analysis, a higher Rad-score (odds ratio [OR], 17.775; $P < .001$), higher histologic grade (grade 2 [OR, 4.200; $P = .026$]; grade 3 [OR, 23.999; $P = .010$]), high Ki-67 expression (OR, 8.093; $P < .001$), and p53 overexpression (OR, 12.280; $P = .003$) were associated with non-low OD risk (Table 2). We considered 2 multivariate logistic regression models. For a clinicopathologic model, a multivariate logistic regression with stepwise selection was performed for clinicopathologic factors, and Ki-67 (OR, 4.878; $P = .015$) and p53 (OR, 4.972; $P = .086$) were selected as predictors. For the combined model, we included Rad-score in the multivariate logistic regression model and then performed stepwise selection. The Rad-score and Ki-67, and p53 overexpression were selected from the procedure and a higher Rad-score (OR, 65.2092; $P < .001$), high Ki-67 expression (OR, 17.462; $P = .007$), and high p53 (OR = 8.449; $P = .077$) were associated with non-low OD risk (Table 3).

ROC curve analysis using a logistic regression model, with Rad-score-only model, yielded an AUC of 0.759. The clinico-pathologic model, with Ki-67 and p53 as the predictors, yielded

Table 2

Univariate logistic regression analyses of the variables associated with predicting non-low Oncotype DX risk.

Variable	Odds ratio	Standard Error	P value
Rad-score	17.775	0.760	<.001
Age, yr	0.990	0.034	.779
Tumor type			
IDC	Ref	Ref	
ILC	0.000	2284.102	.994
Mucinous carcinoma	0.000	3956.180	.997
Tumor size, cm	0.888	0.361	.742
Histologic grade			
1	Ref	Ref	
2	4.200	0.643	.026
3	23.999	1.242	.010
Extensive intraductal component			
Absent	Ref	Ref	
Present	0.680	0.571	.499
Lymphovascular invasion			
Absent	Ref	Ref	
Present	0.846	0.554	.762
PR			
Negative	Ref	Ref	
Positive	0.000	1978.090	.993
Ki-67			
<14%	Ref	Ref	
≥14%	8.093	0.590	<.001
p53			
Negative	Ref	Ref	
Positive	12.280	0.848	.003
T stage			
1	Ref	Ref	
2	1.313	0.598	.649
N stage			
0	Ref	Ref	
1	1.029	0.622	.963
Tumor shape			

(continued)

Table 3

(continued).

Variable	Odds ratio	Standard Error	P value
Oval or round	Ref	Ref	
Irregular	0.506	0.528	.197
Tumor margin			
Circumscribed	Ref	Ref	
Not circumscribed	0.554	0.527	.262
Internal enhancement			
Homogeneous	Ref	Ref	
Heterogeneous	0.457	0.838	.350
Rim	4.572	0.776	.050
Enhancement kinetics of delayed phase			
Persistent	Ref	Ref	
Plateau	11513145.184	1455.398	.991
Washout	2576648.290	1455.398	.992
Peak enhancement value, %	1.000	0.002	.832
Percentage of washout, %	0.993	0.012	.541

HER2 = human epidermal growth factor 2, IDC = invasive ductal carcinoma, ILC = invasive lobular carcinoma, PR = progesterone receptor.

Table 3

Multivariate logistic regression analyses with stepwise selection of variables associated with predicting non-low Oncotype DX risk.

Model/Variable	Odds ratio	Standard Error	P value
Clinicopathologic model			
Ki-67	4.878	0.653	.015
p53	4.972	0.934	.086
Combined model			
Rad-score	65.209	1.137	<.001
Ki-67	17.462	1.072	.007
p53	8.449	1.207	.077

an AUC of 0.574. For the combined model, the AUC increased to 0.900 (Fig. 2). The optimal cut-off for the Rad-score model was 0.154. Figures 3 and 4 show the representative cases in the low and non-low risk groups along with corresponding radiomics feature values.

4. Discussion

In this study, we investigated a quantitative radiomics signature for ER-positive invasive breast cancers on 3T DCE-MRI and assessed whether there was an association with OD RS. Multivariate analyses revealed that a higher Rad-score (OR, 65.209; $P<.001$), high Ki-67 expression (OR, 17.462; $P=.007$), and high p53 (OR = 8.449; $P=.077$) are associated with non-low OD risk. ROC curve analyses showed that the risk classification performance of the Rad-score-only model (AUC = 0.759) was better than the clinicopathological model (AUC = 0.574), and the best performance was achieved when the Rad-score was added to clinicopathological model (AUC = 0.900).

The OD currently represents the most commonly used multigene assay that has been validated for both the prediction of chemotherapy treatment benefit and the estimation of the risk of distant recurrence in patients with ER-positive, early-stage breast cancer.[1,31] Paik et al[3] in the National Surgical Adjuvant
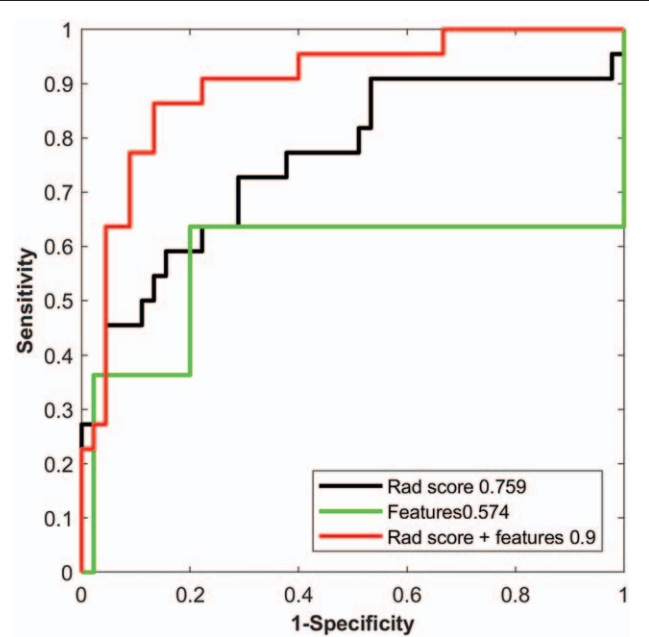

Figure 2. Receiver operating characteristic curves and area under the receiver operating characteristic values from 3 different models: a) Rad-score-only model, b) clinicopathologic model, and c) combined model.

Breast and Bowel Project (NSABP) B-20 trial reported that patients with high RS tumors (>30) had a large benefit from chemotherapy and patients with low RS (<18) derived minimal benefit from chemotherapy. Similar findings have been reported for node-positive patients. In the Southwest Oncology Group (SWOG) 8814 trial, Albain et al[8] showed that RS was strong predictive factor of chemotherapy benefit for 10-year disease-free survival, but there was no apparent benefit for scores <18 or 18 to 30. In addition to the prediction of chemotherapy benefit, Paik et al[2] in the NSABP B-14 trial reported that rate of 10-year distant recurrence in the low-risk group (RS <18) was significantly lower in comparison with the high-risk group (RS >30) in the patients with node-negative, tamoxifen-treated breast cancer. The Arimidex, Tamoxifen, Alone or in Combination (ATAC) study by Dowsett et al[9] demonstrated that RS was significantly associated with time to distant recurrence for both node-negative and node-positive patients.

As an effort to predict OD RS by using MRI, several studies have identified radiomics MRI features associated with OD assay. Ashraf et al[21] utilized 4 intrinsic imaging phenotypes extracted with 31 morphologic, kinetic, and spatial heterogeneity 1.5T DCE-MRI features, and explored their correlation with OD; phenotypes 1 and 2 consisted entirely of tumors with an RS <31 and included only low- and intermediate- risk tumors. Sutton et al[15] reported similar results while investigating 44 morphologic and texture imaging DCE-MRI features at 1.5T or 3.0T; multivariate analyses identified 2 texture features (kurtosis on the first [$P=.0056$] and third [$P=.0005$] post-contrast sequences) that significantly correlated with recurrence scores. A correlation study by Wan et al[32] showed that various computer-extracted texture DCE-MRI features at 1.5T were highly correlated with high and low OD risk classifications for ER-positive breast cancers, although they only included patients with low (<18) and high (>30) OD RS, excluding intermediate RS (18–30).

High OD (RS 34) : Rad-Score 0.4245
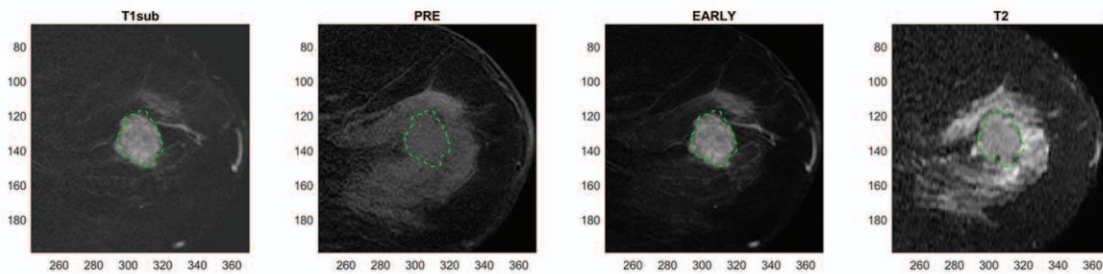
Min_val_T2	21	Energy_val_Sub T1	-0.3396
Skewness_T2	0.2624	Max_probability_val_Sub T1	0.0018
Percentile_histogram.2.5_T2	37	Skewness_PRE T1	-0.5385
Contrast_glcm_val_T2	307.2401	Sum_variance_val_PRE T1	0
Skewness_Sub T1	1	Variance_glcm_val_CE T1	957.3958

Figure 3. The representative MRI images in the non-low risk group along with corresponding radiomics feature values in a 58-year-old woman with invasive ductal carcinoma. Her OD RS was 34 and categorized as non-low risk group. Rad-score based on radiomic features was 0.42, which was concordant with the category assessed by OD RS. MRI = magnetic resonance imaging, OD = Oncotype DX, RS = recurrence score.

Low OD (RS 12): Rad-Score -0.1989
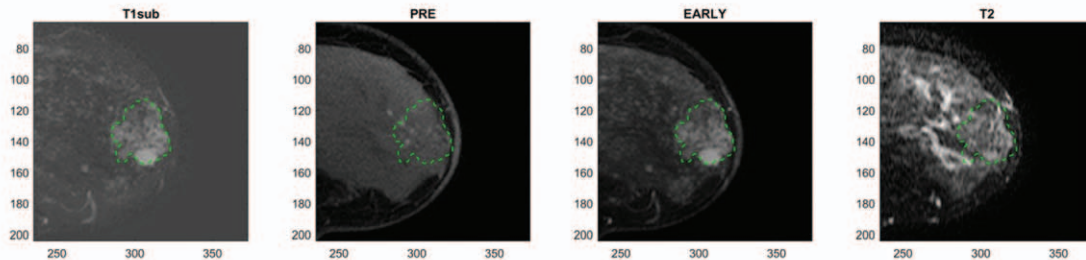
Min_val_T2	1	Energy_val_Sub T1	0.0416
Skewness_T2	0.5460	Max_probability_val_Sub T1	0.0010
Percentile_histogram.2.5_T2	10	Skewness_PRE T1	0.1771
Contrast_glcm_val_T2	335.9048	Sum_variance_val_PRE T1	0
Skewness_Sub T1	1	Variance_glcm_val_CE T1	484.9705

Figure 4. The representative MRI images in the low risk group along with corresponding radiomics feature values in a 33-year-old woman with invasive ductal carcinoma. Her OD RS was 12 and categorized as low risk group. Rad-score based on radiomic features was -0.20, which was concordant with the category assessed by OD RS. MRI = magnetic resonance imaging, OD = Oncotype DX, RS = recurrence score.

However, these previous studies used relatively limited radiomics features on small patient cohorts using different machines.

For Rad-score construction, 156 candidate radiomics features were reduced to 10 potential predictors by using the LASSO method. The combined analyses of a panel of multiple predictors as a radiomics signature has been embraced in recent studies[24–26] to estimate disease-free survival in early-stage non-small cell lung cancer[24] and predict lymph node metastasis in colorectal cancer.[26] Although the AUC (0.900) of the combined model was still incomplete to replace OD RS, the Rad-score showed the potential for a noninvasive, fast, economic, and repeatable method for use during treatment in routine practice, which could replace more expensive genomic testing. To the best of our knowledge, this is the first attempt to investigate the association between OD RS classifications and a radiomics signature, as well as various clinicopathologic and conventional MRI findings, including CAD data.

Among the established histopathologic parameters, high Ki-67 expression (OR, 17.462; $P = .007$) remained a significant factor associated with the non-low OD risk group. Our results are

consistent with other research. Williams et al[12] showed that Ki-67 was significantly associated with OD RS, Nottingham grade, and angiolymphatic invasion. Sahebjam et al[33] reported a strong correlation between Ki-67 and OD RS, especially in tumors with a Ki-67 value ≥25%. Moreover, the likelihood of a tumor with a Ki-67 value ≥25% with an intermediate or high OD RS was >90%, and such patients may benefit from adjuvant chemotherapy.

Our study has several limitations. First, our results may be hard to generalize. We retrospectively reviewed MR images at a single tertiary academic institution using a single machine and MR protocol in limited, homogeneous condition. Second, the number of intermediate or high OD risk patients was relatively small compared to the number of low-risk patients. However, considering OD assay was usually performed in patients expected to be low risk, this was inevitable. Also, in clinical practice, patients with intermediate risk usually undergo adjuvant chemotherapy. Therefore, we divided the study population into low (RS <18) and non-low (RS ≥18) OD risk groups for statistical robustness and clinical reality. The number of non-low OD risk patients (22/67, 32.8%) was not small. Third, large number of extracted radiomics features and relatively small sample size may lead to model overfitting limiting the generalizability of the results. Additional validation study is needed. Fourth, we did not perform intra- or inter-observer reproducibility of radiomics feature extraction or Rad-score. However, Huang et al[26] analyzed inter-observer and intra-observer reproducibility of radiomics feature extraction using inter- and intra-class correlation coefficients (ICCs), and reported good agreement (ICC >0.75). Fifth, the clinical significance of our results might be underestimated as our methodology for manual ROI drawing was time-consuming and labor-intensive. There are automatic and semi-automatic ROI segmentation algorithms that do not suffer from disadvantages of manual approach, which is left for future work. Our research shows possibilities for using a radiomics signature for risk stratification as a proof of concept study. In addition, our volumetric measurement could reflect the true tumor heterogeneity instead of single slice. Finally, we constructed the Rad-scores using linear combination of the selected features and non-linear method of could offer better performance. Such non-linear models require more samples than we currently have and thus is left for future work.

In conclusion, this study revealed that the Rad-score incorporating radiomics features of DCE-MRI is a significant predictor for discriminating between low and non-low OD risk groups in patients with ER-positive invasive breast cancers. The radiomics signature (Rad-score) showed the potential for non-invasive prediction of which ER-positive patients might have little-to-no benefit from adjuvant chemotherapy.

Author contributions

Conceptualization: Eun Sook Ko.

Data curation: Hyunjin Park, Eun Sook Ko, Yaeji Lim, Hwan-Ho Cho.

Formal analysis: Kyung Jin Nam, Hyunjin Park, Yaeji Lim, Hwan-Ho Cho.

Investigation: Kyung Jin Nam, Hyunjin Park, Eun Sook Ko, Yaeji Lim, Hwan-Ho Cho.

Methodology: Hyunjin Park, Eun Sook Ko, Yaeji Lim, Hwan-Ho Cho, Jeong Eon Lee.

Project administration: Hwan-Ho Cho.

Resources: Kyung Jin Nam, Hwan-Ho Cho.

Software: Hyunjin Park, Hwan-Ho Cho.

Supervision: Eun Sook Ko, Yaeji Lim, Jeong Eon Lee.

Validation: Yaeji Lim.

Visualization: Yaeji Lim.

Writing – original draft: Kyung Jin Nam, Hyunjin Park, Eun Sook Ko.

Writing – review & editing: Hyunjin Park, Eun Sook Ko, Yaeji Lim, Hwan-Ho Cho, Jeong Eon Lee.

References

[1] Markopoulos C. Overview of the use of Oncotype DX (®) as an additional treatment decision tool in early breast cancer. Expert Rev Anticancer Ther 2013;13:179–94.

[2] Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351:2817–26.

[3] Paik S, Tang G, Shak S, 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.

[4] Mamounas EP, Tang G, Fisher B, et al. Association between the 21-gene recurrence score assay and risk of locoregional recurrence in node-negative, estrogen receptor-positive breast cancer: results from NSABP B-14 and NSABP B-20. J Clin Oncol 2010;28:1677–83.

[5] Tang G, Cuzick J, Costantino JP, et al. Risk of recurrence and chemotherapy benefit for patients with node-negative, estrogen receptor-positive breast cancer: recurrence score alone and integrated with pathologic and clinical factors. J Clin Oncol 2011;29:4365–72.

[6] Tang G, Shak S, Paik S, et al. Comparison of the prognostic and predictive utilities of the 21-gene Recurrence Score assay and Adjuvant! for women with node-negative, ER-positive breast cancer: results from NSABP B-14 and NSABP B-20. Breast Cancer Res Treat 2011;127:133–42.

[7] Habel LA, Shak S, Jacobs MK, et al. A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. Breast Cancer Res 2006;8:R25.

[8] Albain KS, Barlow WE, Shak S, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 2010;11:55–65.

[9] Dowsett M, Cuzick J, Wale C, et al. Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study. J Clin Oncol 2010;28:1829–34.

[10] Harris LN, Ismaila N, McShane LM, et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American society of clinical oncology clinical practice guideline. J Clin Oncol 2016;34:1134–50.

[11] Basavanhally A, Feldman M, Shih N, et al. Multi-field-of-view strategy for image-based outcome prediction of multi-parametric estrogen receptor-positive breast cancer histopathology: comparison to Oncotype DX. J Pathol Inform 2011;2:S1.

[12] Williams DJ, Cohen C, Darrow M, et al. Proliferation (Ki-67 and phosphohistone H3) and oncotype DX recurrence score in estrogen receptor-positive breast cancer. Appl Immunohistochem Mol Morphol 2011;19:431–6.

[13] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6.

[14] Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. Magn Reson Imaging 2012;30:1234–48.

[15] Sutton EJ, Oh JH, Dashevsky BZ, et al. Breast cancer subtype intertumor heterogeneity: MRI-based features predict results of a genomic assay. J Magn Reson Imaging 2015;42:1398–406.

[16] Mazurowski MA. Radiogenomics: what it is and why it is important. J Am Coll Radiol 2015;12:862–6.

[17] Zinn PO, Mahajan B, Sathyan P, et al. Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. PLoS One 2011;6:e25451.

[18] Gevaert O, Xu J, Hoang CD, et al. Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data–methods and preliminary results. Radiology 2012;264: 387–96.

[19] Yamamoto S, Maki DD, Korn RL, et al. Radiogenomic analysis of breast cancer using MRI: a preliminary study to define the landscape. AJR Am J Roentgenol 2012;199:654–63.

[20] Mazurowski MA, Zhang J, Grimm LJ, et al. Radiogenomic analysis of breast cancer: luminal B molecular subtype is associated with enhancement dynamics at MR imaging. Radiology 2014;273:365–72.

[21] Ashraf AB, Daye D, Gavenonis S, et al. Identification of intrinsic imaging phenotypes for breast cancer tumors: preliminary associations with gene expression profiles. Radiology 2014;272:374–84.

[22] Birkhahn M, Mitra AP, Cote RJ. Molecular markers for bladder cancer: the road to a multimarker approach. Expert Rev Anticancer Ther 2007;7:1717–27.

[23] Croner RS, Förtsch T, Brückl WM, et al. Molecular signature for lymphatic metastasis in colorectal carcinomas. Ann Surg 2008;247: 803–10.

[24] Huang Y, Liu Z, He L, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. Radiology 2016;281:947–57.

[25] Liang C, Huang Y, He L, et al. The development and validation of a CT-based radiomics signature for the preoperative discrimination of stage I-II and stage III-IV colorectal cancer. Oncotarget 2016;7:31401–12.

[26] Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol 2016;34:2157–64.

[27] Meyer CR, Boes JL, Kim B, et al. Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. Med Image Anal 1997;1:195–206.

[28] Morris EA, Comstock CE, Lee CH, et al. ACR BI-RADS magnetic resonance imaging. In: ACR BI-RADS Atlas, breast imaging reporting and data system. Reston, Va: American College of Radiology, 2013.

[29] Kadam P, Bhalerao S. Sample size calculation. Int J Ayurveda Res 2010;1:55–7.

[30] van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017;77:e104–7.

[31] Simon R. Development and validation of biomarker classifiers for treatment selection. J Stat Plan Inference 2008;138:308–20.

[32] Wan T, Bloch BN, Plecha D, et al. A radio-genomics approach for identifying high risk estrogen receptor-positive breast cancers on DCE-MRI: preliminary results in predicting OncotypeDX Risk Scores. Sci Rep 2016;6:21394.

[33] Sahebjam S, Aloyz R, Pilavdzic D, et al. Ki 67 is a major, but not the sole determinant of Oncotype Dx recurrence score. Br J Cancer 2011; 105:1342–5.