Clinical Cancer Research

Radiomics Signature on Magnetic Resonance Imaging: Association with Disease-Free Survival in Patients with Invasive Breast Cancer

Hyunjin Park[1,2], Yaeji Lim[3], Eun Sook Ko[4], Hwan-ho Cho[5], Jeong Eon Lee[6], Boo-Kyung Han[4], Eun Young Ko[4], Ji Soo Choi[4], and Ko Woon Park[4]

Abstract

Purpose: To develop a radiomics signature based on preoperative MRI to estimate disease-free survival (DFS) in patients with invasive breast cancer and to establish a radiomics nomogram that incorporates the radiomics signature and MRI and clinicopathological findings.

Experimental Design: We identified 294 patients with invasive breast cancer who underwent preoperative MRI. Patients were randomly divided into training ($n = 194$) and validation ($n = 100$) sets. A radiomics signature (Rad-score) was generated using an elastic net in the training set, and the cutoff point of the radiomics signature to divide the patients into high- and low-risk groups was determined using receiver-operating characteristic curve analysis. Univariate and multivariate Cox proportional hazards model and Kaplan–Meier analysis were used to determine the association of the radiomics signature, MRI findings, and clinicopathological variables with DFS. A radiomics nomogram combining the Rad-score and MRI and clinicopathological findings was constructed to validate the radiomic signatures for individualized DFS estimation.

Results: Higher Rad-scores were significantly associated with worse DFS in both the training and validation sets ($P = 0.002$ and 0.036, respectively). The radiomics nomogram estimated DFS [C-index, 0.76; 95% confidence interval (CI); 0.74–0.77] better than the clinicopathological (C-index, 0.72; 95% CI, 0.70–0.74) or Rad-score–only nomograms (C-index, 0.67; 95% CI, 0.65–0.69).

Conclusions: The radiomics signature is an independent biomarker for the estimation of DFS in patients with invasive breast cancer. Combining the radiomics nomogram improved individualized DFS estimation. *Clin Cancer Res; 24(19); 4705–14.* *©2018 AACR.*

Introduction

As breast cancers are spatially and temporally heterogeneous, typical invasive biopsy methods do not accurately provide a complete characterization of the entire tumor. Imaging has great potential in guiding therapy because it can provide a more comprehensive view of the entire tumor and can be used repeatedly during treatment to monitor disease and/or the response to therapy.

Radiomics is an emerging methodology that automatically extracts high dimensional features from imaging data, which can later be mined and analyzed for decision support (1–4). Recent advances in radiomics have provided insights in personalized medicine in oncologic practice related to tumor detection, subtype classification, and assessment of treatment response (3, 5–9). A set of relevant biomarkers, commonly referred to as signature information, has been defined; it has been shown to have promising results, and it may be powerful enough to change clinical management (10–13).

Many studies involving survival analysis of patients with tumors of various organs have reported that several texture features, such as uniformity and entropy, can be used in risk stratification using a limited number of texture features (1, 7, 9, 14–18). Because prognosis is not determined by a single risk factor, these previously described approaches have a limited value because they use only few selected features. Therefore, multifeature signatures are deemed reasonable, and huge numbers of markers are usually studied in a high-throughput manner (10, 12, 13, 19, 20). Although it is now recognized that a signature composed of multiple biomarkers holds higher value than a single biomarker (10), to the best of our knowledge, a radiomics signature associated with recurrence in patients with invasive breast cancer has not yet been reported. If radiomics features from preoperative staging MRI can be used to predict patient outcomes, they can be particularly beneficial because MRI is noninvasive and lacks additional cost.

Therefore, the aim of this study was to develop a radiomics signature based on preoperative MRI to estimate the disease-free

[1]School of Electronic and Electrical Engineering, Sungkyunkwan University, Jangan-gu, Suwon, Korea. [2]Center for Neuroscience Imaging Research, Institute for Basic Science (IBS), Jangan-gu, Suwon, Korea. [3]Department of Applied Statistics, Chung-Ang University, Dongjak-gu, Seoul, Korea. [4]Department of Radiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Gangnam-gu, Seoul, Korea. [5]Department of Electronic, Electrical and Computer Engineering, Sungkyunkwan University, Jangan-gu, Suwon, Korea. [6]Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Gangnam-gu, Seoul, Korea.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

H. Park and Y. Lim contributed equally to the article.

Corresponding Author: Eun Sook Ko, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Korea. Phone: 82-2-3410-0877; Fax: 82-2-3410-0049; E-mail: mathilda0330@gmail.com

doi: 10.1158/1078-0432.CCR-17-3783

Translational Relevance

Because of the tumor heterogeneity, typical invasive biopsy methods do not accurately provide a complete characterization of the entire tumor. Radiomics is an emerging methodology that automatically extracts high-dimensional features from imaging data, which facilitates the use of these imaging features to quantify differences between tissues that are imperceptible to human eyes. In this study, the identified radiomics signature can be used as a biomarker for risk stratification for disease-free survival (DFS) in patients with invasive breast cancer. In addition, a radiomics nomogram that incorporates the radiomics signature, MRI findings, and clinicopathological findings, can be used to facilitate the preoperative individualized prediction of recurrence in patients with breast cancer. Such quantitative radiomics prognostic models of breast cancer may potentially be useful in precision medicine, and they can affect patients' treatment strategies.

survival (DFS) in patients with invasive breast cancer and to further establish a radiomics nomogram that incorporates the radiomics signature, MRI findings, and clinicopathological findings for the individual preoperative prediction of recurrence.

Materials and Methods

Patients

Institutional review board approval was obtained for this study (SMC 2017-08-136) and the need for informed patient consent was waived owing to its retrospective nature. This study was conducted in accordance with the Declaration of Helsinki.

Eight-hundred twenty-three consecutive women (mean age, 50 years) who underwent surgery for invasive breast cancer between March 2011 and December 2011 were identified. The inclusion criteria of our study were as follows: (i) preoperative dynamic contrast-enhanced MRI using a 1.5-T scanner at our institution; (ii) initial unilateral breast malignancy with a final pathologic diagnosis of invasive breast cancer; and (iii) a lesion presenting as a mass on MRI. The exclusion criteria were as follows: (i) MRI performed in patients after the diagnosis of cancer by vacuum-assisted or excisional biopsy; (ii) patients treated with neoadjuvant chemotherapy (NAC); (iii) patients presenting with metastatic disease and/or a concurrent malignancy; (iv) patients with positive margins of resection. Finally, 294 cancers in 294 women (mean age, 51 years; range, 24–85 years) were included in this study.

MRI protocol

All patients underwent dynamic contrast-enhanced (DCE) MRI. Details regarding the acquisition parameters and MRI retrieval procedure are presented in Appendix E1 (online).

MRI 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 texture 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. Details regarding the MRI preparation for radiomics analysis are presented in Appendix E2 (online). To calculate the interobserver agreement of feature extraction, we randomly selected 49 patients using statistical software and again placed the regions-of-interest (ROI) on contrast-enhanced T1-weighted subtraction images; this, was performed by a different radiologist with 6 years of experience in breast MRI (J.S.C). The interclass correlation coefficient (ICC) was calculated and an ICC of greater than 0.75 was considered to represent good agreement.

Conventional MRI and clinicopathological evaluations

The MR images of the masses were retrospectively evaluated according to the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) MR lexicon (21) by two board-certified radiologists (J.S.C. and E.S.K., with 6 and 10 years of experience in breast MRI, respectively) in consensus. Kinetic features on computer-aided diagnosis (CAD) were also recorded. The CAD reports included the following kinetic features used in our analysis: the initial peak enhancement values; proportions of early-phase medium and rapid enhancements; and proportion of delayed-phase persistent, plateau, and washout enhancements. Details regarding the interpretation of MRI are presented in Appendix E3 (online).

The final histopathological results of surgical specimens were reviewed to determine the following: tumor size; histological grade; presence of extensive intraductal component (EIC); presence of lymphovascular invasion; estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2); and Ki-67 expression status. Tumors with HER2 scores of 3+ (strong homogeneous staining) were considered positive. In the case of 2+ scores (moderate complete membranous staining in $\geq$1% of tumor cells), silver in situ hybridization (SISH) was used to determine HER2 amplification. Breast cancers were divided into three molecular subtypes based on the immunohistochemical or SISH findings for ER, PR, and HER2 as follows: luminal (hormone receptor–positive and any HER2 status), HER2-enriched (hormone receptor-negative and HER2-positive), and triple-negative (hormonal receptor–negative and HER2-negative) (22). For convenience, pathologic diagnoses were divided into three groups: invasive ductal carcinoma, invasive lobular carcinoma, and others.

The end point of our study was DFS, which was defined as the time from the date of surgery to that of the first recurrence of the disease, date of death, date last known to have no evidence of disease, or date of the most recent follow-up. Disease recurrence was defined as the outcome of breast cancer recurrence (local-regional or distant) or new primary contralateral breast cancer (invasive or ductal carcinoma in situ). Information regarding the patient follow-up and recurrence status was obtained from patient medical records or from clinicians involved in the treatment and follow-up. Data concerning patients who did not have recurrence at the last follow-up or who were lost to follow-up were treated as censored in the analysis.

Statistical analysis

A two-tailed P value of <0.05 was considered as statistically significant. All statistical analyses were performed by a dedicated statistician using R statistical software (version 3.2.4; R Foundation for Statistical Computing, Vienna, Austria). We used the "glmnet" package to perform the elastic net Cox regression model analysis. Kaplan-Meier curve, nomogram construction and calibration plot

Table 1. Characteristics of patients in the training set and validation set

Characteristics	Training set ($n = 194$)	Validation set ($n = 100$)	*P*
Age, y[a]	50.64 ± 10.46	52.34 ± 10.51	0.190
Tumor size (mm)[a]	21.19 ± 11.77	19.98 ± 12.05	0.408
Rad-score[a]	0.00 ± 0.06	0.00 ± 0.07	0.696
Early: Peak enhancement (%)[a]	256.09 ± 109.87	259.65 ± 92.28	0.782
Early: Medium (%)[a]	8.07 ± 17.72	6.47 ± 15.93	0.448
Early: Rapid (%)[a]	91.96 ± 17.73	93.56 ± 15.95	0.450
Delayed: Persistent (%)[a]	21.00 ± 22.76	19.67 ± 24.47	0.646
Delayed: Plateau (%)[a]	42.22 ± 18.32	38.05 ± 17.60	0.062
Delayed: Washout (%)[a]	36.78 ± 26.88	42.28 ± 26.91	0.097
T stage			0.064
1	111 (57.22%)	70 (70%)	
2	77 (39.69%)	26 (26%)	
3	5 (2.58%)	4 (4%)	
4	1 (0.52%)	0 (0%)	
N stage			0.516
0	126 (64.95%)	58 (58%)	
1	50 (25.77%)	34 (34%)	
2	12 (6.19%)	6 (6%)	
3	6 (3.09%)	2 (2%)	
Stage			0.550
1	84 (43.30%)	50 (50%)	
2	88 (45.36%)	40 (40%)	
3	22 (11.34%)	10 (10%)	
Pathologic type			0.038
IDC	171 (88.14%)	95 (95%)	
ILC	11 (5.67%)	0 (0%)	
Others	12 (6.19%)	5 (5%)	
Mass shape			1.000
Round/oval	49 (25.26%)	25 (25%)	
Irregular	145 (74.74%)	75 (75%)	
Mass margin			0.682
Circumscribed	17 (8.76%)	11 (11%)	
Not circumscribed	177 (97.24%)	89 (89%)	
Internal enhancement			0.232
Homogeneous	18 (9.28%)	16 (16%)	
Heterogeneous	148 (76.29%)	71 (71%)	
Rim enhancement	28 (14.43%)	13 (13%)	
Histologic grade			0.933
1	57 (29.38%)	28 (28%)	
2	81 (41.75%)	44 (44%)	
3	56 (28.89%)	28 (28%)	
ER			0.673
Positive	138 (71.13%)	68 (68%)	
Negative	56 (28.87%)	32 (32%)	
PR			0.098
Positive	131 (67.53%)	57 (57%)	
Negative	63 (32.47%)	43 (43%)	
HER2			0.158
Positive	37 (19.07%)	27 (27%)	
Negative	157 (80.93%)	73 (73%)	
Molecular subtype			0.263
Luminal	140 (72.16%)	68 (68%)	
HER2-enriched	17 (8.76%)	15 (15%)	
Triple-negative	37 (19.07%)	17 (17%)	
Ki-67			0.568
≥14%	120 (61.86%)	66 (66%)	
<14%	74 (38.14%)	34 (34%)	
Lymphovascular invasion			0.965
Present	64 (32.99%)	34 (34%)	
Absent	130 (67.01%)	66 (66%)	
Extensive intraductal component			0.595
Present	38 (19.59%)	23 (23%)	
Absent	156 (80.41%)	77 (77%)	
Adjuvant chemotherapy			0.704
No	52 (26.80%)	24 (24%)	
Yes	142 (73.20%)	76 (76%)	
Adjuvant radiotherapy			0.965
No	29 (14.95%)	14 (14%)	
Yes	165 (85.05%)	86 (86%)	
Adjuvant endocrine therapy			0.742
No	57 (29.38%)	32 (32%)	
Yes	137 (70.62%)	68 (68%)	

NOTE: Unless otherwise noted, data are numbers of patients, with percentages in parentheses.
[a]Data are means ± standard deviations.

were analyzed using the "rms" and "hdnom" packages. Note that some R functions were modified to apply to the data.

Radiomics analysis, Rad-score building, and validation of Rad-score

Radiomics features were computed volumetrically over the radiologist-drawn ROIs from each of the four distinct MRI series. The features were computed using a combination of open source code (23) and in-house generated computer code implemented in MATLAB (Mathworks Inc.). Details of computer code used to compute the radiomics features are described and provided in Appendix E4. A total of 156 texture features in three distinct categories were computed. The features were grouped into morphological (eight features), histogram-based (19 features), and higher-order texture features (18 features). In Appendix E5 and Appendix Table S1, the categorical concepts of radiomics features (Appendix E5) and the mathematical definition of the adopted feature algorithms (Appendix Table 1) are described (online). Tumor volume was also assessed based on the ROIs drawn by the study radiologist.

We randomly divided patients into two groups, a training set ($n = 194$) and validation set ($n = 100$) using statistical software. Patients' characteristics in the training set and validation set were compared using an analysis of variance (ANOVA) for continuous variables and a chi-squared test or Fisher's exact test for categorical variables. The elastic net method, which is known to be ideal for the regression of high-dimensional data and also useful in variable selection for highly correlated variables (24), was used to select the most useful predictive radiomics features from the training set. A radiomics signature (Rad-score) was calculated for each patient via a linear combination of selected features that were weighted by their respective coefficients. ICC value of the Rad-score that was determined by two radiologists was calculated.

The potential association of the radiomics signature with DFS was first assessed in the training set and then validated in the validation set. The patients were classified into high-risk or low-risk groups according to the Rad-score, the threshold of which was identified using receiver-operating characteristic (ROC) curve analysis. We employed the Youden's J statistics (25), which uses the maximum value (sensitivities + specificities) as the optimality criterion. Kaplan-Meier curves were used to analyze survival between the high- and low-risk groups. Log-rank tests were used to compare differences in the survival between the two groups. In the training set, the univariate Cox proportional hazards model was used to analyze the effects of clinicopathological variables (age, adjuvant chemotherapy, adjuvant radiation therapy, adjuvant endocrine therapy, histological grade, T stage, N stage, EIC, lymphovascular invasion, molecular subtype, and Ki-67 status), morphologic and quantitative volumetric factors obtained via

Table 2. Characteristics of patients according to the risk group based on radiomics signature in the training set

	Low-risk ($n = 84$)	High-risk ($n = 109$)	P
Age, y[a]	50.43 ± 8.88	50.68 ± 11.53	0.869
Tumor size (mm)[a]	16.48 ± 10.67	24.86 ± 11.35	<0.001
Rad-score[a]	−0.05 ± 0.04	0.05 ± 0.03	<0.001
Early: Peak enhancement (%)[a]	221.67 ± 80.97	282.39 ± 122.09	<0.001
Early: Medium (%)[a]	10.23 ± 21.14	6.49 ± 14.53	0.147
Early: Rapid (%)[a]	89.82 ± 21.14	93.54 ± 14.55	0.15
Delayed: Persistent (%)[a]	23.60 ± 25.50	19.17 ± 20.35	0.181
Delayed: Plateau (%)[a]	44.10 ± 21.08	40.69 ± 15.88	0.202
Delayed: Washout (%)[a]	32.30 ± 27.70	40.12 ± 25.95	0.045
T stage			<0.001
1	67 (79.76%)	43 (39.45%)	
2	15 (17.86%)	62 (56.88%)	
3	2 (2.38%)	3 (2.75%)	
4	0 (0%)	1 (0.92%)	
N stage			0.003
0	66 (78.57%)	59 (54.13%)	
1	14 (16.67%)	36 (33.03%)	
2	2 (2.38%)	10 (9.17%)	
3	2 (2.38%)	4 (3.67%)	
Stage			<0.001
1	56 (66.67%)	27 (24.77%)	
2	24 (28.57%)	64 (58.72%)	
3	4 (4.76%)	18 (16.51%)	
Pathologic type			0.009
IDC	68 (80.95%)	102 (93.58%)	
ILC	6 (7.14%)	5 (4.59%)	
Others	10 (11.90%)	2 (1.83%)	
Mass shape			0.582
Round/oval	24 (28.57%)	24 (22.02%)	
Irregular	60 (71.43%)	85 (77.98%)	
Mass margin			0.123
Circumscribed	4 (4.76%)	13 (11.93%)	
Not circumscribed	80 (95.24%)	96 (88.07%)	
Internal enhancement			0.087
Homogeneous	9 (10.71%)	8 (7.34%)	
Heterogeneous	68 (80.95%)	80 (73.39%)	
Rim enhancement	7 (8.33%)	21 (19.27%)	
Histologic grade			0.003
1	37 (44.05%)	19 (17.43%)	
2	29 (34.52%)	52 (47.71%)	
3	18 (21.43%)	38 (34.86%)	
ER			<0.001
Positive	66 (78.57%)	71 (65.14%)	
Negative	18 (21.43%)	38 (34.86%)	
PR			0.002
Positive	62 (73.81%)	68 (62.39%)	
Negative	22 (26.19%)	41 (37.61%)	
HER2			0.274
Positive	13 (15.48%)	24 (22.02%)	
Negative	71 (84.52%)	85 (77.98%)	
Ki-67			0.07
≥14%	42 (50.00%)	78 (71.56%)	
<14%	42 (50.00%)	31 (28.44%)	
Lymphovascular invasion			0.041
Present	20 (23.81%)	44 (40.37%)	
Absent	64 (76.19%)	65 (59.63%)	
Extensive intraductal component			0.590
Present	15 (17.86%)	23 (21.10%)	
Absent	69 (82.14%)	86 (78.90%)	

NOTE: Unless otherwise noted, data are numbers of patients, with percentages in parentheses.

[a]Data are means ± standard deviations. Rad-score could not be calculated in one patient who was excluded.

MRI (mass shape, mass margin, internal enhancement pattern, and percentage volume of each kinetic component found within the tumor at early and delayed phases of enhancement), and Rad-score on DFS. Variables significant in the univariate Cox proportional hazard model ($P < 0.05$) were included in the multivariate Cox proportional hazard model. To overcome the multicollinearity, we then performed stepwise selection based on the Akaike information criterion (AIC).

Development and validation of the radiomics nomogram

To demonstrate the value of the radiomics signature, the radiomics nomogram, Rad-score-only nomogram, and clinicopathological nomogram were evaluated in the training set and then validated in the validation set. The radiomics nomogram incorporated the radiomics signature and independent various risk factors based on multivariate Cox analysis with stepwise selection. The performance of the radiomics nomogram was compared with that of both the Rad-score-only and clinicopathological nomograms. The clinicopathological nomogram did not include the Rad-score or variables associated with MRI. An index of probability of concordance (C-index) between the predicted probability and actual outcome was calculated to evaluate the predictive ability and discrimination of the model (26). The value of the C-index ranges from 0.5 to 1.0, with 0.5 indicating random chance and 1.0 indicating a perfectly accurate discrimination. The nomograms were subjected to bootstrapping validation (1,000 bootstrap resamples) to calculate a relatively corrected C-index. To improve statistical robustness of results, we divided patients and performed the same procedures three additional times. The mean C-index of the radiomics nomogram and Rad-score-only nomogram in the additional divisions was calculated.

Results

Patient outcomes

Two-hundred thirty-nine (239/294, 81.3%) patients underwent breast-conserving surgery and mastectomy was performed in 55 (55/294, 18.7%) patients. Regarding survival outcomes, there were 32 recurrences (13 local-regional, five contralateral breast, and 14 distant recurrences) after a mean follow-up period of 54.2 months (range, 5–64 months). The mean time to recurrence was 33.1 months (range, 5–60 months). There was one patient with a recurrence within the first 6 months of follow-up, and this patient might have been due to residual disease.

Patients' characteristics in the training set and validation set are shown in Table 1. Except for the pathologic type, no differences were found between the training and validation sets in Rad-score or clinicopathological findings. The mean tumor volume calculated was 4,168.9 mm^3 (range, 72.7–7,2876.0). Other pathologic diagnoses in the 17 patients were as follows: mucinous carcinoma (7), medullary carcinoma (2), invasive micropapillary carcinoma (3), tubular carcinoma (1), invasive apocrine carcinoma (3), and adenoid cystic carcinoma (1).

The interobserver agreement of the radiomics features extraction between two readers was extremely high. The mean ICC value for the 156 radiomics features was 0.960 (range, 0.760–0.999).

Creation of the radiomics signature-based prediction model

The 156 texture features were divided into four groups [one morphological group and three higher-order texture groups (gray

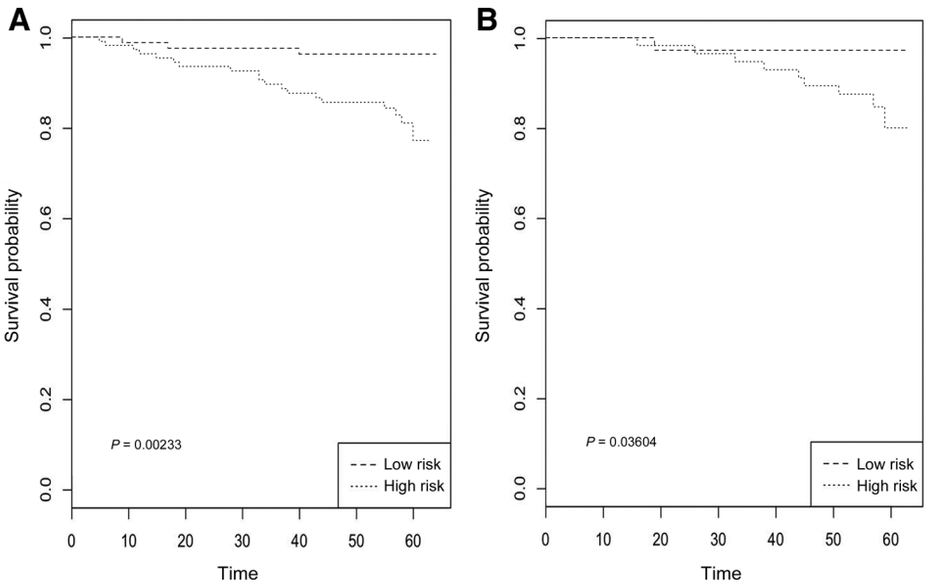

Figure 1.
Kaplan–Meier survival analyses according to the radiomics signature for patients in the training set (**A**) and those in the validation set (**B**). A significant association of the radiomic signature with the DFS was shown in the training set, which was then confirmed in the validation set.

level co-occurrence matrix, GLCM)] based on the potential predictors determined from the 194 patients in the training set and from the features with nonzero coefficients in the elastic net Cox proportional regression model. Because the variables had diverse ranges of values, we standardized the texture values for the elastic net regression. These features are described in the Rad-score calculation formula in the Appendix E6 (online). The mean ICC value of the Rad-score was 0.937 [95% confidence interval (CI), 0.891–0.964].

The median Rad-score was 0.0134 (range, −0.2179 to 0.0951; interquartile range, −0.0342 to 0.0517). The optimum cutoff value generated by the ROC curve was -0.003 (95% CI, 0.474–0.864). Using this threshold value, patients were classified into a high-risk group (Rad-score ≥ -0.003) and a low-risk group (Rad-score < -0.003). The characteristics of the patients according to risk group are shown in Table 2. In the training set, larger tumor size ($P < 0.001$), higher Rad-score ($P < 0.001$), higher peak enhancement ($P < 0.001$), higher washout component ($P = 0.045$), higher T stage ($P < 0.001$), higher N stage ($P = 0.003$), pathologic diagnosis of IDC ($P = 0.009$), higher histological grade ($P = 0.003$), ER negativity ($P < 0.001$), PR negativity ($P = 0.002$), and presence of lymphovascular invasion ($P = 0.041$) were associated with the high-risk group. Kaplan–Meier curves showed that the radiomics signature was associated with the DFS in the training set ($P = 0.002$), and this finding was confirmed in the validation set ($P = 0.036$; Fig. 1).

Results of the univariate analysis are shown in Table 3. A higher Rad-score, larger tumor size, higher peak enhancement on CAD, T2 stage, N3 stage, higher histological grade, ER negativity, PR negativity, triple-negative subtype, and not receiving adjuvant endocrine therapy were associated with worse DFS. In the multivariate analysis with stepwise selection, higher Rad-score (DFS hazard ratio, 256300.000; 95% CI, 18.315–3585000000; $P = 0.011$), N3 stage (DFS hazard ratio, 5.269; 95% CI, 1.621–17.130; $P = 0.006$) and triple-negative subtype (DFS hazard ratio, 7.461; 95% CI, 2.922–19.050; $P < 0.0001$) remained independent prognostic factors in the Cox proportional hazards model (Table 4).

Performance and validation of the radiomics nomogram

A nomogram that incorporated the significant factors was established (Fig. 2). The calibration curve of the radiomics nomogram for estimating DFS outcome demonstrated good agreement between prediction and observation in the validation set (Fig. 3A). The C-index of the nomogram for the prediction of poorer recurrence outcome was 0.76 (95% CI, 0.74–0.77). The calibration curve of a clinicopathological nomogram that included tumor size, N stage and molecular subtype is also shown (Fig. 3B). The C-index of the nomogram for the prediction of poorer recurrence outcome was 0.72 (95% CI, 0.70–0.74). Figure 3C shows the calibration curve of the Rad-score–only nomogram, which had a C-index of 0.64 (95% CI, 0.63–0.69). The mean C-index of the additional three divisions was 0.76 (95% CI, 0.75–0.78) in radiomics nomogram and 0.58 (95% CI, 0.56–0.60) in the Rad-score–only nomogram.

Discussion

Medical imaging is one of the major tools in oncologic diagnosis, treatment guidance, and treatment monitoring. Unlike biopsy, imaging can be used noninvasively to assess the characteristics of human tissue, which is why it is routinely used in clinical practice. However, traditional imaging interpretation is subjective or qualitative, which limits its clinical significance. Recent advances in medical imaging acquisition and analysis, including radiomics, allow high-throughput extraction of informative imaging features that can quantify the differences between tissues that are imperceptible to human eyes.

Texture analysis has been used in breast MRI to detect microcalcifications (27), differentiate between benign and malignant lesions (28, 29), and distinguish between breast cancer subtypes (16, 18). In addition, this method has been used to predict treatment response in patients treated with NAC (7). Concerning the relationship between outcomes in patients treated with NAC and texture features, Pickles and colleagues showed that higher entropy on DCE-MR images was associated with poorer outcomes in patients treated with NAC (8). In a preoperative setting, Kim

Table 3. Univariate analysis of DFS in the training set

Variable	Hazard ratio (95% CI)	*P*
Age	0.997 (0.957–1.039)	0.894
Tumor size	1.033 (1.004–1.064)	0.026
Early: Peak enhancement (%)	1.003 (1.000–1.006)	0.023
Early: Medium (%)	0.989 (0.958–1.020)	0.468
Early: Rapid (%)	1.011 (0.981–1.043)	0.471
Delayed: Persistent (%)	0.992 (0.970–1.013)	0.443
Delayed: Plateau (%)	0.997 (0.974–1.019)	0.767
Delayed: Washout (%)	1.007 (0.991–1.022)	0.401
T stage		
1	Ref	
2	2.434 (1.008–5.876)	0.048
3	3.141 (0.393–25.139)	0.281
4	0 (0–∞)	0.998
N stage		
0	Ref	
1	1.370 (0.506–3.706)	0.536
2	0.911 (0.117–7.071)	0.929
3	9.519 (3.009–30.114)	<0.001
Stage		
1	Ref	
2	2.352 (0.829–6.678)	0.108
3	4.022 (1.162–13.922)	0.028
Pathologic type		
IDC	Ref	
ILC	0.784 (0.105–5.846)	0.813
Others	0.648 (0.087–4.836)	0.672
Mass shape		
Round/oval	Ref	
Irregular	1.756 (0.591–5.216)	0.311
Mass margin		
Circumscribed	Ref	
Not circumscribed	1.021 (0.238–4.370)	0.978
Internal enhancement		
Homogeneous	Ref	
Heterogeneous	1.049 (0.241–4.564)	0.949
Rim enhancement	1.639 (0.299–8.971)	0.569
Histologic grade		
1	Ref	
2	2.487 (0.517–11.975)	0.256
3	7.819 (1.763–34.675)	0.007
ER		
Positive	Ref	
Negative	5.182 (2.171–12.368)	0.0002
PR		
Positive	Ref	
Negative	4.086 (1.713–9.745)	0.002
HER2		
Positive	Ref	
Negative	2.427 (0.567–10.383)	0.232
Molecular subtype		
Luminal	Ref	
HER2-enriched	1.006 (0.123–7.945)	0.995
Triple-negative	6.213 (2.612–14.776)	<0.0001
Ki-67		
≥14%	Ref	
<14%	0.329 (0.111–0.973)	0.044
Lymphovascular invasion		
Present	Ref	
Absent	0.384 (0.166–0.890)	0.026
Extensive intraductal component		
Present	Ref	
Absent	1.133 (0.383–3.349)	0.821
Adjuvant chemotherapy		
Yes	Ref	
No	3.884 (0.907–16.621)	0.067
Adjuvant radiotherapy		
Yes	Ref	
No	1.441 (0.487–4.255)	0.509
Adjuvant endocrine therapy		
Yes	Ref	
No	4.167 (1.776–9.804)	0.001
Rad-score	35656.46 (6.182–205645034)	0.018

(Continued in the following column)

and colleagues evaluated the relationship between MRI texture features and survival outcomes in 203 patients with primary breast cancer (9). Almost all of these prior studies have typically used few radiomics features. This methodology seems anecdotal and intuitive and might lead to an underestimation of the significance of radiomics. Similar problems have occurred in the field of radiogenomics, and construction of multifactor panels is a more common approach to overcome this challenge in outcome estimation (30, 31).

In this background, results of several studies attempting to build models by using numerous radiomics features to predict the molecular subtype have been recently published (16, 32). Waugh and colleagues (16) used texture analysis from 220 imaging features to find the surrogate markers of the molecular subtypes in 148 cancers and 73 test sets. In a study by Grimm and colleagues (32), 56 imaging features, including morphologic, texture, and dynamic features, were evaluated as surrogate markers of the molecular subtypes in 275 breast cancers. This approach has also been used to investigate the correlation between the 21-gene recurrence score (Oncotype DX, Genomic Health) or other kinds of multigene assays, such as those of MammaPrint or PAM50, and breast MRI findings (33–36). For example, in a study by Sutton and colleagues (35), 44 morphologic and texture features were investigated in 95 patients who also underwent 21-gene recurrence score testing and preoperative breast MRI. On multivariate analysis, kurtosis on the first ($P = 0.0056$) and third ($P = 0.0005$) post-contrast sequences was significantly correlated with recurrence scores.

Taking a step forward, several researchers have developed and validated a radiomics signature-based nomogram for the preoperative prediction of lymph node metastasis in colorectal cancer (12) or for the prediction of DFS in early-stage non–small cell lung cancer (13). In this study, we developed and validated a radiomics signature-based nomogram for the preoperative individualized prediction of recurrence in patients with breast cancer. The nomogram incorporates three components of the radiomics signature based on 156 features from four image series, MR kinetic features assessed using a commercially available CAD system, and clinicopathological findings. The nomogram can facilitate preoperative individualized prediction of poor outcome with respect to DFS.

Table 4. Multivariate analysis of DFS in the training set

Variable	Hazard ratio (95% CI)	*P*
Rad-score	256300.000 (18.315–3585000000)	0.011
N stage		
0	Ref	
1	1.635 (0.575–4.652)	0.357
2	0.348 (0.043–2.827)	0.323
3	5.269 (1.621–17.130)	0.006
Molecular subtype		
Luminal	Ref	
HER2-enriched	0.844 (0.104–6.837)	0.873
Triple-negative	7.461 (2.922–19.050)	<0.0001

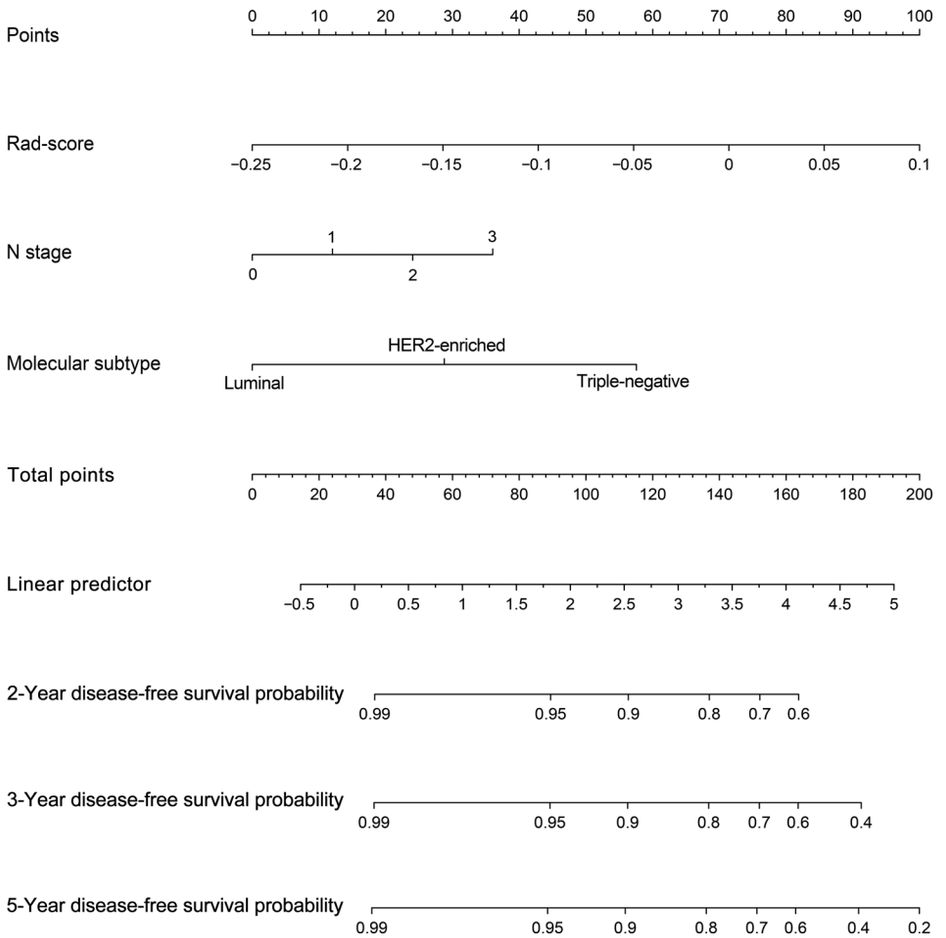

Figure 2.
The developed radiomics nomogram. The radiomics nomogram was developed using the training set using the radiomics score. The Rad-score was determined by drawing a vertical line to the points' axis to determine how many points toward the probability of DFS the patient receives. The process was repeated for each variable, and the points for each of the risk factors were added. The final total was then located on the Total Point axis.

Interestingly, in our study, a total of four radiomics features were selected (i.e., surface to volume ratio [SVR], cluster tendency of gray level co-occurrence matrix [GLCM] from T2, variance of GLCM from T2, and sum variance of GLCM from T2); the well-known radiomics feature, entropy, was not included. Many prior studies have particularly emphasized the significance of entropy (7, 9, 6–18, 29). We suspect that entropy was not selected in the elastic net regression due to multicollinearity because we used numerous texture features that could interact with each other. Briefly, SVR was selected from the shape features. The SVR quantifies the degree of irregularity of the tumor boundary. Irregular tumor boundary could be associated with poor survival. Three GLCM-related features were selected, which all of them reflect textural heterogeneity within the tumor. Cluster tendency of GLCM measures the homogeneity of the textural patterns. The sum variance of GLCM measures the dispersion of the textural information using the sum of grayscale values. The variance of GLCM is proportional to the rate of occurrence of different textural patterns within the tumor. These GLCM features have different mathematical definitions and thus they measure different aspects of tumor textural heterogeneity. Unlike histogram-based features, which are dependent on a single pixel value, texture features, which are based on GLCM/gray level size zone matrix (GLSZM), take into account the interaction between neighboring pixels and are thus, they are well suited to quantify tumor texture and heterogeneity. This is consistent with previous observations that texture features based on GLCM/GLSZMs are better at capturing heterogeneous texture information than histogram-based features (37). In addition, three of the four features selected were calculated from T2-weighted images. This is in accordance with prior speculation that T2-weighted images depict better tumor heterogeneity than contrast-enhanced T1-weighted images do (7).

Our study has three strengths. First, this is the first study (to our knowledge) performed to estimate survival in patients with invasive breast cancer using a radiomics signature. Our study demonstrated that the radiomics signature can be used to predict survival outcomes and a combined radiomics–clinicopathological nomogram achieved superior prognostic performance than either the Rad-score-only or the clinico-pathological nomograms, with a higher C-index and better calibration. Second, our radiomics signature was calculated using ROIs drawn on entire tumors. Most previous studies (12, 13, 33, 35) only used the single largest slice, which does not reflect the true tumor heterogeneity. In addition, the single-slice methodology may raise concerns of selection bias. Third, we used standardized texture values for the elastic net regression because all texture features had diverse ranges. We believe this procedure increased the accuracy of interpretation of the texture features.

Our study has several limitations. First, this study had a relatively small sample size and a relatively short follow-up period. Second,

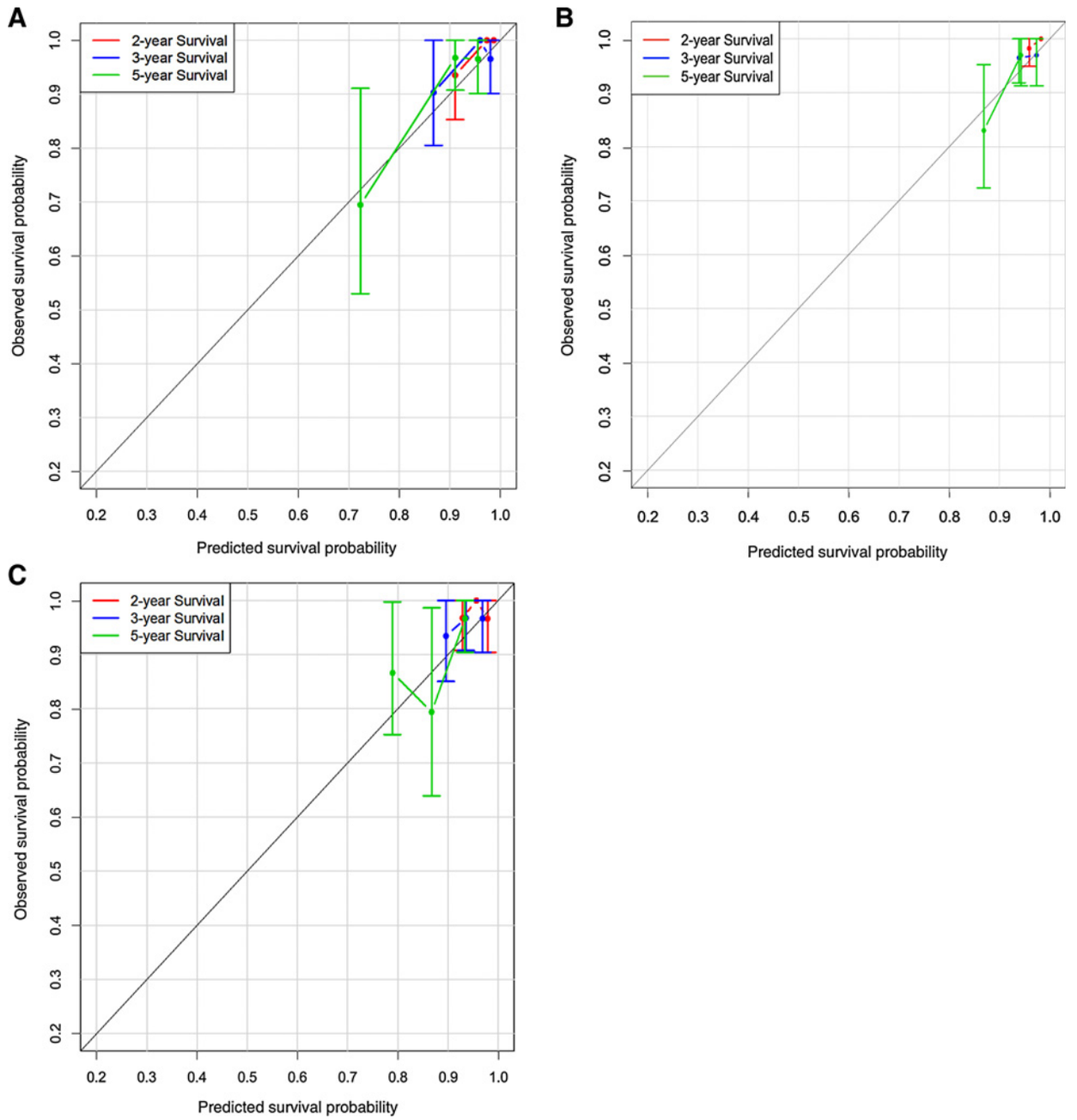

Figure 3.

Calibration curves of three nomograms in the validation set. **A,** Calibration curve of the radiomics nomogram. **B,** Calibration curve of the clinicopathological nomogram. **C,** Calibration curve of the Rad-score–only nomogram. The dashed line represents a perfect match between the nomogram-predicted probability (*x*-axis) and the actual probability calculated using the Kaplan–Meier analysis (*y*-axis).

our results might be hard to generalize because it was performed using a single machine in a single institution with same MRI protocol on mass lesions only. Further studies involving subjects with various conditions in a larger population are needed. Third, we did not conduct external validation to prove the robustness of our results using an independent dataset. To overcome this limitation, we performed 3 additional divisions, and the results were equally good. External validation is left as a future study. Fourth, we did not include other possible risk factors, such as family history or genetic abnormalities. Fifth, placement of the ROI on the whole tumor and calculation of radiomics parameters were not automatically performed, making these time-consuming and labor-intensive tasks. This may weaken the clinical application and significance of our study. However, we believe our study, which was primarily a proof-of-concept study, showed the potential of using radiomics in daily practice. In the near future, we hope that commercially available

software that performs automatic segmentation of the tumor and calculation of radiomics parameters will be developed.

In conclusion, our results showed the identified radiomics signature has the potential to be used as a biomarker for risk stratification for DFS in patients with invasive breast cancer. In addition, our study showed that a radiomics nomogram that incorporates the radiomics signature, MRI findings, and clinicopathological findings, can be used to facilitate the preoperative individualized prediction of recurrence in patients with breast cancer. Such quantitative radiomics prognostic models of breast cancer may potentially be useful for precision medicine and they affect patients' treatment strategies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: H. Park, E.S. Ko
Development of methodology: H. Park, Y. Lim, E.S. Ko, H.-H. Cho
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): E.S. Ko, B.-K. Han, E.Y. Ko, J.S. Choi, K.W. Park
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): H. Park, Y. Lim, H.-H. Cho, J.E. Lee
Writing, review, and/or revision of the manuscript: H. Park, Y. Lim, E.S. Ko, J.E. Lee, J.S. Choi
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): E.S. Ko
Study supervision: E.S. Ko, J.E. Lee

Acknowledgments

We appreciate Ho Yun Lee for her technical assistance. This study was supported by the Institute for Basic Science (grant number, IBS-R015-D1) and the National Research Foundation of Korea (grant number, NRF-2016R1A2B4008545).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 20, 2017; revised April 12, 2018; accepted June 11, 2018; published first June 18, 2018.

References

1. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6.
2. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. Magn Reson Imaging 2012;30:1234–48.
3. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumor phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
4. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563–77.
5. Altazi B, Fernandez D, Zhang G, Biagioli M, Moros E, Lee Moffitt H. SU-E-J-258: prediction of cervical cancer treatment response using radiomics features based on F18-FDG uptake in PET images. Med Phys 2016;42:3326.
6. Li H, Lan L, Drukker K, Perou C, Giger M. TU-AB-BRA-08: radiomics in the analysis of breast cancer heterogeneity on DCE-MRI. Med Phys 2016;42:3588.
7. Parikh J, Selmi M, Charles-Edwards G, Glendenning J, Ganeshan B, Verma H, et al. Changes in primary breast cancer heterogeneity may augment midtreatment MR imaging assessment of response to neoadjuvant chemotherapy. Radiology 2014;272:100–12.
8. Pickles MD, Lowry M, Gibbs P. Pretreatment prognostic value of dynamic contrast- enhanced magnetic resonance imaging vascular, texture, shape, and size parameters compared with traditional survival indicators obtained from locally advanced breast cancer patients. Invest Radiol 2016;51:177–85.
9. Kim JH, Ko ES, Lim Y, Lee KS, Han BK, Ko EY, et al. Breast cancer heterogeneity: MR imaging texture analysis and survival outcomes. Radiology 2017;282:665–75.
10. Birkhahn M, Mitra AP, Cote RJ. Molecular markers for bladder cancer: the road to multimarker approach. Expert Rev Anticancer Ther 2007;7:1717–27.
11. Croner RS, Förtsch T, Brückl WM, Rödel C, Papadopoulos T, Brabletz T, et al. Molecular signature for lymphatic metastasis in colorectal carcinomas. Ann Surg 2008;247:803–10.
12. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, 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.
13. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, 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.
14. Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K. Tumor heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. Eur Radiol 2012;22:796–802.
15. Win T, Miles KA, Janes SM, Ganeshan M, Shastry M, Endozo R, et al. Tumor heterogeneity and permeability as measured on the CT component of PET/CT predict survival in patients with non-small cell lung cancer. Clin Cancer Res 2013;19:3591–9.
16. Waugh SA, Purdie CA, Jordan LB, Vinnicombe S, Lerski RA, Martin P, et al. Magnetic resonance imaging texture analysis classification of primary breast cancer. Eur Radiol 2016;26:322–30.
17. Sinha S, Lucas-Quesada FA, DeBruhl ND, Sayre J, Farria D, Gorczyca DP, et al. Multifeature analysis of Gd-enhanced MR images of breast lesions. J Magn Reson Imaging 1997;7:1016–26.
18. Holli K, Laaperi AL, Harrison L, Luukkaala T, Toivonen T, Ryymin P, et al. Characterisation of breast cancer subtypes by texture analysis of magnetic resonance images. Acad Radiol 2010;17:135–41.
19. Mitra AP, Lam LL, Ghadessi M, Erho N, Vergara IA, Alshalalfa M, et al. Discovery and validation of novel expression signature for postcystectomy recurrence in high-risk bladder cancer. J Natl Cancer Inst 2014;106:dju290.
20. Mitra AP, Pagliarulo V, Yang D, Waldman FM, Datar RH, Skinner DG, et al. Generation of a concise gene panel for outcome prediction in urinary bladder cancer. J Clin Oncol 2009;27:3929–37.
21. American College of Radiology. Breast imaging reporting and data system (BI-RADS). 5th ed: VA: American College of Radiology, 2013.
22. Blaschke E, Abe H. MRI phenotype of breast cancer: kinetic assessment for molecular subtypes. J Magn Reson Imaging 2015;42:920–4.
23. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017;77:e104–7.
24. Zou H, Hastie T. Regularization and Variable Selection via the Elastic Net. J R Stat Soc Series B Stat Methodol 2005;67:301–20.
25. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32–5
26. Wolbers M, Koller MT, Witteman JC, Steyerberg EW. Prognostic models with competing risks: methods and application to coronary risk prediction. Epidemiology 2009;20:555–61.
27. James D, Clymer BD, Schmalbrock P. Texture detection of simulated microcalcification susceptibility effects in magnetic resonance imaging of breasts. J Magn Reson Imaging 2001;13:876–81.
28. Chen W, Giger ML, Li H, Bick U, Newstead GM. Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. Magn Reson Med 2007;58:562–71.
29. Gibbs P, Turnbull LW. Textural analysis of contrast-enhanced MR images of the breast. Magn Reson Med 2003;50:92–8.

30. Rutman AM, Kuo MD. Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. Eur J Radiol 2009;70:232–41.
31. Kuo MD, Gollub J, Sirlin CB, Ooi C, Chen X. Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinomas. J Vasc Interv Radiol 2007;18:821–31.
32. Grimm LJ, Zhang J, Mazurowski MA. Computational approach to radiogenomics of breast cancer: luminal A and luminal B molecular subtypes are associated with imaging features on routine breast MRI extracted using computer vision algorithms. J Magn Reson Imaging 2015;42:902–7.
33. Ashraf AB, Daye D, Gavenonis S, Mies C, Feldman M, Rosen M, et al. Identification of intrinsic imaging phenotypes for breast cancer tumors: preliminary associations with gene expression profiles. Radiology 2014;272:374–84.
34. Dialani V, Gaur S, Mehta TS, Venkataraman S, Fein-Zachary V, Phillips J, et al. Prediction of low versus high recurrence scores in estrogen receptor-positive, lymph node-negative invasive breast cancer on the basis of radiologic-pathologic features: comparison with Oncotype DX test recurrence scores. Radiology 2016;280:370–8.
35. Sutton EJ, Oh JH, Dashevsky BZ, Veeraraghavan H, Apte AP, Thakur SB, et al. Breast cancer subtype intertumor heterogeneity: MRI-based features predict results of a genomic assay. J Magn Reson Imaging 2015;42:1398–406.
36. Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, 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;281:382–91.
37. Davnall F, Yip CS, Ljungqvist G, Selmi N, Ng F, Sanghera B, et al. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? Insights Imaging 2012;3:573–89.