



Machine learning with multiparametric breast MRI for prediction of Ki-67 and histologic grade in early-stage luminal breast cancer

Sung Eun Song¹ · Kyu Ran Cho¹ · Yongwon Cho¹ · Kwangsoo Kim² · Seung Pil Jung³ · Bo Kyoung Seo⁴ · Ok Hee Woo⁵

Received: 30 April 2021 / Revised: 20 May 2021 / Accepted: 9 June 2021 / Published online: 12 August 2021
© European Society of Radiology 2021

Abstract

Objectives To investigate whether machine learning–based prediction models using 3-T multiparametric MRI (mpMRI) can predict Ki-67 and histologic grade in stage I–II luminal cancer.

Methods Between 2013 and 2019, consecutive women with luminal cancers who underwent preoperative MRI with diffusion-weighted imaging (DWI) and surgery were included. For prediction models, morphology, kinetic features using computer-aided diagnosis (CAD), and apparent diffusion coefficient (ADC) at DWI were evaluated by two radiologists. Logistic regression analysis was used to identify mpMRI features for predicting Ki-67 and grade. Diagnostic performance was assessed using eight machine learning algorithms incorporating mpMRI features and compared using the DeLong method.

Results Of 300 women, 203 (67.7%) had low Ki-67 and 97 (32.3%) had high Ki-67; 242 (80.7%) had low grade and 58 (19.3%) had high grade. In multivariate analysis, independent predictors for higher Ki-67 were washout component > 13.5% (odds ratio [OR] = 4.16; $p < 0.001$) and intratumoral high SI on T2-weighted image (OR = 1.89; $p = 0.022$). Those for higher grade were washout component > 15.5% (OR = 7.22; $p < 0.001$), rim enhancement (OR = 2.59; $p = 0.022$), and ADC value < $0.945 \times 10^{-3} \text{ mm}^2/\text{s}$ (OR = 2.47; $p = 0.015$). Among eight models using these predictors, six models showed the equivalent performance for Ki-67 (area under the receiver operating characteristic curve [AUC]: 0.70) and Naive Bayes classifier showed the highest performance for grade (AUC: 0.79).

Conclusions A prediction model incorporating mpMRI features shows good diagnostic performance for predicting Ki-67 and histologic grade in patients with luminal breast cancers.

Key Points

- Among multiparametric MRI features, kinetic feature of washout component > 13.5% and intratumoral high signal intensity on T2-weighted image were associated with higher Ki-67.
- Washout component > 15.5%, rim enhancement, and mean apparent diffusion coefficient value < $0.945 \times 10^{-3} \text{ mm}^2/\text{s}$ were associated with higher histologic grade.
- Machine learning–based prediction models incorporating multiparametric MRI features showed good diagnostic performance for Ki-67 and histologic grade in luminal breast cancers.

Keywords Breast cancer · Magnetic resonance imaging · Ki-67 antigen, diagnosis, computer-assisted · Machine learning

✉ Kyu Ran Cho
krcho@korea.ac.kr

¹ Department of Radiology, Korea University Anam Hospital, Korea University College of Medicine, 73, Goryeodae-ro, Seongbuk-gu, Seoul 02841, Republic of Korea

² Transdisciplinary Department of Medicine and Advanced Technology, Seoul National University Hospital, Seoul, Republic of Korea

³ Department of Surgery, Korea University Anam Hospital, Korea University College of Medicine, Seoul, Republic of Korea

⁴ Department of Radiology, Korea University Ansan Hospital, Korea University College of Medicine, Gyeonggi-do, Republic of Korea

⁵ Department of Radiology, Korea University Guro Hospital, Korea University College of Medicine, Seoul, Republic of Korea

Abbreviations

ADC	Apparent diffusion coefficient
AJCC	American Joint Committee on Cancer
AUC	Area under receiver operating characteristic curve
BI-RADS	Breast Imaging Reporting and Data System
BPE	Background parenchymal enhancement
CAD	Computer-aided diagnosis
DWI	Diffusion-weighted imaging
ER	Estrogen receptor
ET	Endocrine therapy
FGT	Fibroglandular tissue
HER2	Human epidermal growth factor receptor2
HR	Hormone receptor
IHC	Immunohistochemical
LN	Lymph node
mpMRI	Multiparametric MRI
NAC	Neoadjuvant chemotherapy
NME	Non-mass enhancement
OR	Odds ratio
PR	Progesterone receptor
ROC	Receiver operating characteristic
ROI	Region of interest
SI	Signal intensity
T2WI	T2-weighted images

Introduction

The eighth edition of the cancer staging system of the American Joint Committee on Cancer (AJCC), announced in 2017, reported that the most important change in breast cancer is the incorporation of biomarkers (histologic grade, estrogen receptor [ER] or progesterone receptor [PR], human epidermal growth factor receptor2 [HER2] expression), which could affect survival into the anatomic staging to make clinical prognostic stages [1]. Additionally, the eighth edition identifies the division of breast cancer into four molecular subtypes: luminal A (hormone receptor [HR] such as ER- or PR-positive, HER2-negative, low Ki-67), luminal B (HR-positive, HER2-negative, high Ki-67), HER2 (HER2-positive regardless of the HR status), and basal (both HR- and HER2-negative) [1–3].

Luminal cancer accounts for 73.2% of all breast cancers [4]. Subgroups (A and B) are classified based on a cutoff value of 14% Ki-67 indicating tumor proliferation [5, 6]. Luminal A subgroup accounts for 48.4% of all breast cancers and typically presents with pT1 tumors, is lymph node (LN)-negative, and responds best to endocrine therapy (ET). Luminal B subgroup accounts for 24.8% and typically presents more commonly with pT2 tumors and is LN-positive and resistant to ET; thus, these cases benefit from combined chemotherapy and ET [4,

7–9]. Besides Ki-67, histologic grade is another important prognostic factor in luminal cancer, which reflects tumor differentiation. A study reported that patients with histologic grade 1 luminal cancers have a better prognosis than those with histologic grade 3 [10]. Therefore, the AJCC manual combines the histologic grade with anatomic staging to determine the clinical prognostic stage, in which luminal cancers with histologic grade 3 are categorized one stage higher than their anatomic stage [1, 11]. For example, in a patient with an anatomic stage IIA luminal cancer, clinical prognostic stage is changed to stage IIB when the histologic grade is 3; however, this case would remain as stage IIA if the histologic grade is 1 or 2 [1, 11].

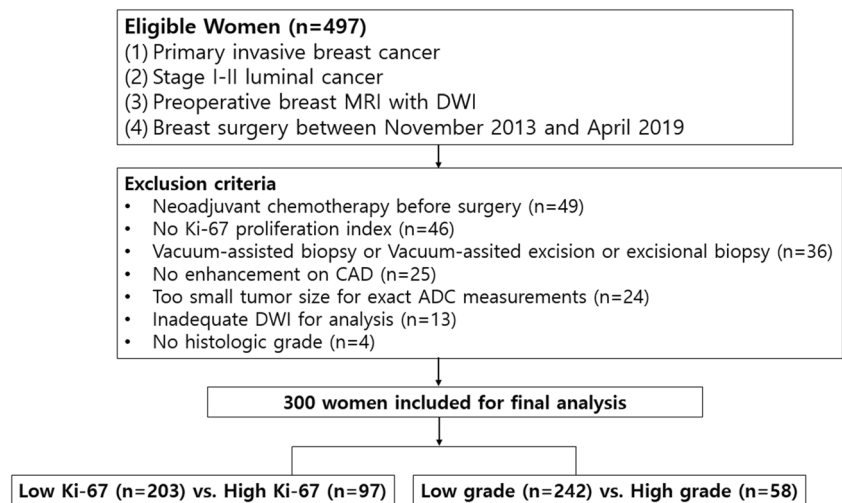
To efficiently select the luminal B cancer patients who may benefit from chemotherapy, many studies have been performed using preoperative MRI features including apparent diffusion coefficient (ADC) at diffusion-weighted imaging (DWI) [12, 13], texture features [14], and imaging features extracted from computer vision algorithms [15]. However, not including partial imaging features but including multiparametric MRI (mpMRI) features is essential for developing elaborate prediction models; 3.0-T mpMRI can provide detailed information about tumors such as morphology, kinetic features related to angiogenesis, and function at the cellular levels via DWI [16]. In the present study, we aimed to investigate whether prediction models based on 8 machine learning algorithms using mpMRI features can predict Ki-67 levels and histologic grade in cases of stage I–II luminal cancer.

Materials and methods

This retrospective study was approved by our institutional review board and the requirement for written informed consent was waived.

Study population

A retrospective review of medical records at our institution between November 2013 and April 2019 was performed, and 497 women were selected consecutively who meet the following criteria: initially diagnosed with invasive breast cancers via a core needle biopsy, confirmed immunohistochemically as luminal cancers (ER- or PR-positive and HER-2 negative), underwent preoperative MRI with DWI followed by surgery, and staged as stage I–II. Among them, 197 patients were excluded, for the following reasons: received neoadjuvant chemotherapy (NAC) before surgery ($n = 49$), no recorded Ki-67 proliferation index ($n = 46$), underwent previous vacuum-assisted biopsy or vacuum-assisted excision or excisional biopsy before MRI examination ($n = 36$), cancer did not have enhancement on CAD and was unable to acquire kinetic features ($n = 25$), tumor size was too small to perform exact ADC

Fig. 1 Flowchart of the study population

measurements on DWI ($n = 24$), DWI was suboptimal for analysis ($n = 13$), and no histologic grade recorded ($n = 4$). Finally, 300 women (mean age \pm standard deviation, 56.6 years \pm 12.0; age range, 32–93 years) were enrolled (Fig. 1). After an initial diagnosis of breast cancer by imaging-guided biopsy, MRI was performed. In all cases, no clip was inserted after biopsy. The interval between biopsy and MRI ranged from 1 to 39 days (median: 13 days).

MRI examination

MRI was performed using a 3.0-T scanner (Achieva 3.0T TX; Philips Healthcare) with a dedicated seven-channel breast array coil in the axial orientation. All scans were taken with the patients in the prone position. Bilateral axial T2-weighted images (T2WI) with fat suppression were acquired (repetition time ms/echo time ms, 5375/65; flip angle, 90°, matrix, 620 \times 303; field of view, 340 \times 340 mm²; section thickness, 3 mm; section gap, 0 mm) and DWI was obtained using an echo-planar-imaging sequence with fat suppression in the axial plane (5417/72; flip angle, 90°, matrix, 96 \times 126; field of view, 320 \times 320 mm²; section thickness, 3 mm; section gap, 0 mm; b values, 0 and 1000 s/mm²). One pre-contrast and five post-contrast dynamic series were obtained using a three-dimensional T1-weighted fast spoiled gradient-echo sequence with fat suppression (5/2; matrix, 436 \times 436; flip angle, 12°; field of view, 340 \times 340 mm²; section thickness, 1 mm; no gap) with intravenous injection of a bolus of 0.1 mmol/kg gadoterate (Dotarem; Guerbet).

MRI image analysis

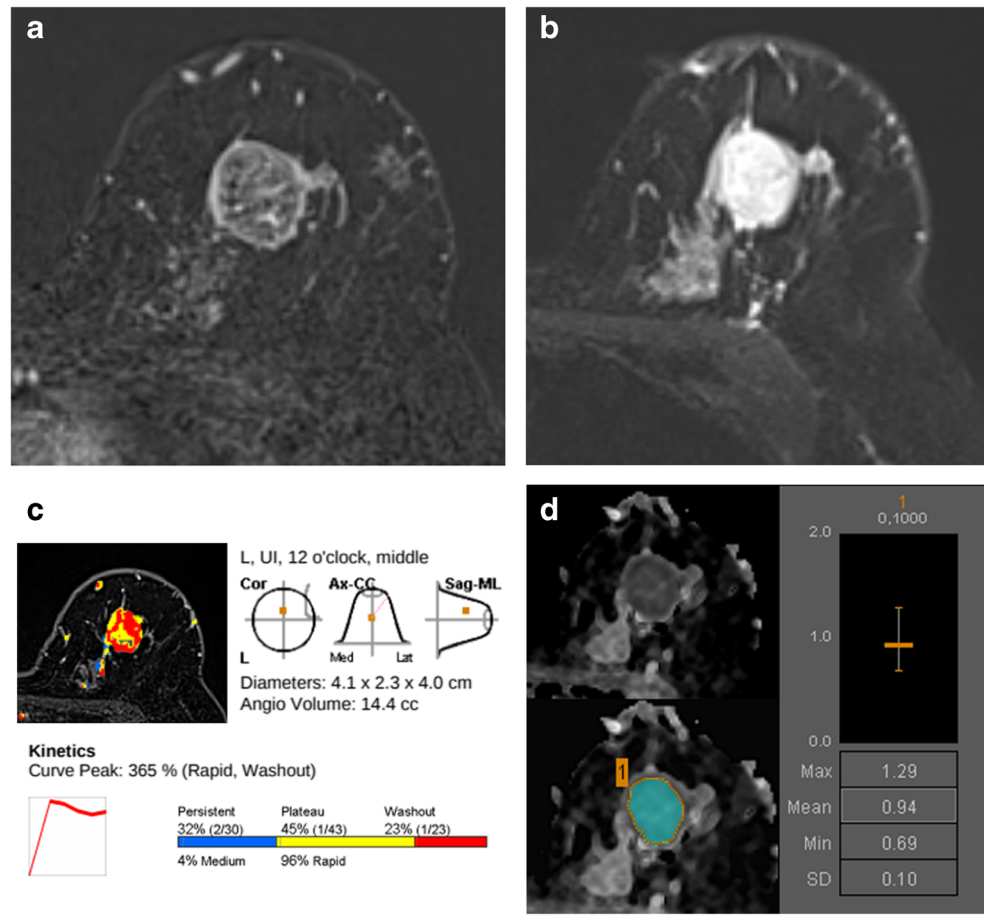
For evaluation of tumor morphology, all MRI images were independently reviewed by two breast radiologists (18 and 9 years of experience), blinded to the patient's clinical history,

using the 2013 Breast Imaging Reporting and Data System MR lexicon [17]. The amounts of fibroglandular tissue (FGT), background parenchymal enhancement (BPE), and the lesion type as mass or non-mass enhancement (NME) were assessed. In patients with multiple lesions, the largest lesion was evaluated. For mass observations, margin, shape, internal enhancement patterns, and associated NME were evaluated. The presence of intratumoral high signal intensity (SI) and peritumoral edema was assessed as absent or present on T2WI [18, 19]. High intratumoral SI on T2WI was visually defined when SI of the lesion was stronger than that of vessels or water or higher than that of the surrounding parenchymal tissue [17–19]. Peritumoral edema was also determined when high SI around the tumor was observed on T2WI [18, 19]. Axillary LN enlargement was defined when one or more following findings were found: abnormal LN shape, increased cortical thickness, irregular LN margins, or completely or partially effaced fatty hilum [17].

Kinetic feature analysis

All T1-weighted images were transferred to a CAD system (CADstream, version 6.0, Confirma) and processed for evaluation of kinetic features. As a threshold of 50–60% has been reported as the most appropriate for the CAD system [20], we used a 50% threshold to classify enhancement. A color overlay angio-map was generated at all enhancing lesions above this threshold. After one radiologist selected the largest tumor on the angio-map, the peak enhancement, angio-volume, and early phase and delayed phase profiles were calculated. Within the tumor, the early enhancement phase profiles were described as either medium or rapid enhancement. The delayed enhancement phase profiles were described as persistent, plateau, or washout.

Fig. 2 Preoperative images of the left breast in a 55-year-old woman with breast cancer. **a** Axial T1-weighted contrast-enhanced subtraction MR image shows rim enhancing mass in the left upper breast. **b** Axial T2-weighted MR image shows intratumoral high signal intensity within the tumor. **c** Auto-portfolio of CAD system shows tumor enhancement kinetics with 23% washout component. **d** Apparent diffusion coefficient (ADC) map shows mass with low signal intensity. A region of interest (green) was manually drawn on the CAD system to encompass an entire cross-section of the lesion. Maximum ADC, mean ADC, and minimum ADC were $1.29 \times 10^{-3} \text{ mm}^2/\text{s}$, $0.94 \times 10^{-3} \text{ mm}^2/\text{s}$, and $0.69 \times 10^{-3} \text{ mm}^2/\text{s}$, respectively. Surgical histopathologic examination revealed a 3-cm invasive ductal carcinoma with a histologic grade 3 and Ki-67 proliferation index of 37% (stage II luminal B cancer)



DWI analysis

All DWI images were transferred to a CAD system and the ADC maps were constructed using b -values of 0 and 1000 s/mm^2 . On the ADC map, two radiologists independently and manually outlined the regions of interest (ROIs) for each lesion. By using T2WI and angio-map as references, measurements were performed twice to cover the entire tumor in its largest cross-sections and to avoid cystic areas (Fig. 2). The minimum, mean, and maximum ADC values of each ROI were measured. The average value of the first ADC measurements by two reviewers was used for statistical analysis.

Histopathologic data analysis

Histopathologic data were determined based on the surgically excised tissue. Using the labeled streptavidin-biotin method for immunohistochemical (IHC) technique, ER and PR positivity were defined as the presence of positive staining in at least 1% of the nuclei in ten high-power fields [21]. HER2 negativity was defined as an IHC score of 0 or 1+ or 2+ staining with negative HER2 gene amplification on fluorescence in situ hybridization. Ki-67 expression was graded as

low ($< 14\%$) or high ($\geq 14\%$) [5, 6] and histologic grade was dichotomized as low (grade 1 or 2) or high (grade 3).

Statistical analysis

The MRI morphologic features, kinetic features, and ADC values were compared between cancers with low Ki-67 and high Ki-67 and in those with low histologic grade and high grade. For categorical variables, the χ^2 test or Fisher exact test was used. For continuous variables, a Student t test was used for normally distributed variables and the Mann-Whitney U test was used for non-normally distributed variables. To determine the optimal cutoff values of kinetic features and ADC parameters for predicting high Ki-67 and histologic grade, receiver operating characteristic (ROC) curve analysis was performed by using the maximum Youden index (i.e., sensitivity + specificity - 1).

To identify the mpMRI features associated with high Ki-67 and histologic grade, logistic regression analysis was used. For multivariate analysis, we used covariates with p values < 0.05 in the univariate analysis. To build a prediction model based on significant mpMRI features from the univariate and multivariate analysis, eight machine learning algorithms (random forest, adaptive boosting, decision tree, k-nearest neighbor, linear

support vector machine, naive Bayes classifier, linear discriminant analysis, and logistic regression) were used. To train the eight machine learning algorithms, repeated ($n = 20$) and 10-fold cross-validation was applied. The DeLong method was used to compare the AUCs of eight algorithms in independent data sets.

Interobserver reliability between two reviewers was assessed using interclass correlation coefficient. An r value

of 1.0 was deemed to indicate perfect agreement; 0.81–0.99, almost perfect agreement; 0.61–0.80, substantial agreement; 0.41–0.60, moderate agreement; 0.21–0.40, fair agreement; and ≤ 0.20 , slight agreement [22]. Statistical analyses were performed using SPSS for Windows, version 20.0 (SPSS Inc.), open-source R software (version 3.5.1; R Foundation for Statistical Computing), and Python (Python Software

Table 1 Associations between multiparametric MRI features and Ki-67 and histologic grade

Features	All ($n = 300$)	Low Ki-67 ($n = 203$)	High Ki-67 ($n = 97$)	p value	Low histologic grade ($n = 242$)	High histologic grade ($n = 58$)	p value
Patient age	56.6 ± 12.0	56.7 ± 12.0	56.4 ± 12.0	0.851	56.3 ± 11.9	57.9 ± 12.5	0.384
Morphologic features							
Amount of fibroglandular tissue				0.856			0.138
Nondense	160	109 (53.7)	51 (52.6)		124 (51.2)	36 (62.1)	
Dense	140	94 (46.3)	46 (47.4)		118 (48.8)	22 (37.9)	
Background parenchymal enhancement				0.884			0.493
Minimal or mild	206	118 (80.3)	117 (79.6)		164 (67.8)	42 (72.4)	
Moderate or marked	94	29 (19.7)	30 (20.4)		78 (32.2)	16 (27.6)	
Lesion type				0.781			0.546
Mass*	285	192 (94.6)	93 (96.6)		229 (94.6)	56 (96.6)	
Non-mass enhancement	15	11 (5.4)	4 (4.1)		13 (5.4)	2 (3.4)	
Mass shape*				0.892			0.430
Round to oval	115	78 (40.6)	37 (39.8)		95 (41.5)	20 (35.7)	
Irregular	170	114 (59.4)	56 (60.2)		134 (58.5)	36 (64.3)	
Mass margin*				0.511			0.652
Circumscribed	55	35 (18.2)	20 (21.5)		43 (18.8)	12 (21.4)	
Not circumscribed	230	157 (81.8)	73 (78.5)		186 (81.2)	44 (78.6)	
Mass internal enhancement*				0.437			0.007
Homo- or heterogeneous	234	160 (83.3)	74 (79.6)		195 (85.2)	39 (69.6)	
Rim	51	32 (16.7)	19 (20.4)		34 (14.8)	17 (30.4)	
Associated non-mass enhancement				0.383			0.987
Absent	202	140 (69.0)	62 (63.9)		163 (67.4)	39 (67.2)	
Present	98	63 (31.0)	35 (36.1)		79 (32.6)	19 (32.8)	
Intratumoral high SI on T2WI				0.001			0.003
No	150	115 (56.7)	35 (36.1)		131 (54.1)	19 (32.8)	
Yes	150	88 (43.3)	62 (63.9)		111 (45.9)	39 (67.2)	
Peritumoral edema on T2WI				0.172			0.127
Absent	248	172 (84.7)	76 (78.4)		204 (84.3)	44 (75.9)	
Present	52	31 (15.3)	21 (21.6)		38 (15.7)	14 (24.1)	
Axillary lymph node enlargement				0.981			0.007
Absent	260	176 (86.7)	84 (86.6)		216 (89.3)	44 (75.9)	
Present	40	27 (13.3)	13 (13.4)		26 (10.7)	14 (24.1)	
Kinetic features on CAD							
Tumor size (cm) †	1.89 ± 0.96	1.79 ± 0.95	2.08 ± 0.95	0.015	1.78 ± 0.90	2.32 ± 1.07	< 0.001
Angio-volume (cm ³) †	2.24 ± 3.40	1.83 ± 2.25	3.09 ± 4.94	0.019	1.79 ± 2.49	2.32 ± 5.49	< 0.001
Peak enhancement (%) †	246.11 ± 130.26	239.59 ± 121.91	259.75 ± 145.94	0.267	238.00 ± 126.86	279.00 ± 139.98	0.057
Early phase-medium component (%) †	36.94 ± 35.18	37.88 ± 35.22	34.96 ± 35.20	0.347	36.93 ± 34.84	36.96 ± 36.90	0.813
Early phase-rapid component (%) †	62.73 ± 35.40	61.63 ± 35.52	65.03 ± 35.21	0.302	62.66 ± 35.11	63.03 ± 36.90	0.760
Delayed phase-persistent component (%) †	49.69 ± 27.88	52.27 ± 28.54	44.30 ± 25.76	0.020	52.98 ± 28.57	35.96 ± 19.70	< 0.001
Delayed phase-plateau component (%) †	32.05 ± 18.12	31.83 ± 19.50	35.50 ± 14.91	0.408	31.35 ± 18.87	34.96 ± 14.33	0.111
Delayed phase-washout component (%) †	18.48 ± 20.40	16.00 ± 20.41	23.68 ± 19.48	0.002	15.82 ± 19.85	29.60 ± 19.02	< 0.001
ADC values on DWI							
Mean ADC †	0.980 ± 0.275	0.988 ± 0.268	0.963 ± 0.289	0.246	0.999 ± 0.281	0.897 ± 0.234	0.011
(× 10 ⁻³ mm ² /s)							
Minimum ADC †	0.597 ± 0.299	0.587 ± 0.303	0.618 ± 0.290	0.542	0.611 ± 0.298	0.540 ± 0.300	0.102
(× 10 ⁻³ mm ² /s)							
Maximum ADC †	1.454 ± 0.344	1.460 ± 0.346	1.440 ± 0.342	0.799	1.457 ± 0.352	1.438 ± 0.314	0.791
(× 10 ⁻³ mm ² /s)							

Data are numbers of patients, with percentages in parentheses

*Mass shape, margin, and internal enhancement were calculated with a denominator of 285 masses

† Numbers are means ± standard deviations

T2WI T2-weighted image, CAD computer-aided diagnosis, ADC apparent diffusion coefficient, DWI diffusion-weighted imaging

Table 2 Univariate and multivariate logistic regression analysis of features associated with high Ki-67

Features	Univariate analysis		Multivariate analysis	
	Odds ratio	<i>p</i> value	Adjusted odds ratio	<i>p</i> value
Intratumoral high SI on T2WI				
Yes	2.31 (1.40 – 3.81)	< 0.001	1.89 (1.11 – 3.20)	0.018
No	Reference		Reference	
Tumor size on CAD				
> 1.75 cm	1.78 (1.09 – 2.91)	0.020		
≤ 1.75 cm	Reference			
Angio-volume on CAD				
> 1.35 cm ³	1.60 (0.98 – 2.61)	0.056		
≤ 1.35 cm ³	Reference			
Persistent component on CAD				
< 37.5 %	1.58 (1.04 – 2.79)	0.032		
≥ 37.5 %	Reference			
Washout component on CAD				
> 13.5 %	3.68 (2.21 – 6.13)	< 0.001	4.16 (2.08 – 8.33)	< 0.001
≤ 13.5 %	Reference		Reference	

Numbers in parentheses are 95% confidence intervals

T2WI T2-weighted image, CAD computer-aided diagnosis, ADC apparent diffusion coefficient

Foundation, version 3.7.4) with Scikit-learn package (<https://github.com/scikit-learn/scikit-learn>).

Results

Patient characteristics and multiparametric MRI features

Of the 300 women, 227 (75.6%) had undergone breast conservation surgery and 73 (24.4%) had undergone total mastectomy. The median pathologic invasive tumor size was 1.5 cm (range, 0.6–4.7 cm) and axillary LN metastasis was confirmed in 25.6% (77 of 300). The range of Ki-67 within this group spanned from 0.1 to 56.6% (median, 12.0%); 203 (67.7%) had a Ki-67 of less than 14%; and 97 (32.3%) had a Ki-67 of 14% or higher. Histologic grade 1 was observed in 91 patients, grade 2 in 151 patients, and grade 3 in 58 patients, and 242 (80.7%) were histologic grade 1 or 2 while 58 (19.3%) were histologic grade 3.

Compared to luminal cancers with low Ki-67, those with high Ki-67 more frequently demonstrated intratumoral high SI on T2WI, larger tumor size, larger angio-volume, smaller persistent component, and larger washout component (all *p* < .05). Compared to luminal cancers with low histologic grade, those with high histologic grade more frequently

Table 3 Univariate and multivariate logistic regression analysis of features associated with high histologic grade

Features	Univariate analysis		Multivariate analysis	
	Odds ratio	<i>p</i> value	Adjusted odds ratio	<i>p</i> value
Intratumoral high SI on T2WI				
Yes	2.42 (1.32 – 4.43)	0.004		
No	Reference			
Mass internal enhancement				
Rim	2.51 (1.27 – 4.94)	0.008	2.59 (1.14 – 5.81)	0.022
Homo- or heterogeneous	Reference		Reference	
Axillary lymph node enlargement				
Absent	2.64 (1.27 – 5.46)	0.009		
Present	Reference			
Tumor size on CAD				
> 1.85 cm	2.86 (1.58 – 5.17)	< 0.001		
≤ 1.85 cm	Reference			
Angio-volume on CAD				
> 1.45 cm ³	2.68 (1.47 – 4.85)	0.001		
≤ 1.45 cm ³	Reference			
Persistent component on CAD				
< 33.5 %	2.43 (1.35 – 4.35)	0.003		
≥ 33.5 %	Reference			
Washout component on CAD				
> 15.5 %	8.21 (4.11 – 16.38)	< 0.001	7.22 (2.97 – 17.57)	< 0.001
≤ 15.5 %	Reference		Reference	
Mean ADC on DWI				
< 0.945 × 10 ⁻³ mm ² /s	2.42 (1.32 – 4.43)	0.004	2.47 (1.19 – 5.13)	0.015
≥ 0.945 × 10 ⁻³ mm ² /s	Reference		Reference	

Numbers in parentheses are 95% confidence intervals

CAD computer-aided diagnosis, ADC apparent diffusion coefficient

demonstrated rim enhancement, intratumoral high SI on T2WI, axillary LN enlargement, larger tumor size, larger angio-volume, smaller persistent component, larger washout component, and lower ADC values (all, $p < .05$) (Table 1).

The optimal cutoff values of kinetic and ADC parameters

Based on the results of the ROC curve analysis, the optimal cutoff values to assess association between high Ki-67 and (1) tumor size was 1.75 cm (area under the curve [AUC], 0.60; $p = .005$); (2) angio-volume, 1.35 cm³ (AUC, 0.58; $p = .019$); (3) persistent component, 37.5% (AUC, 0.58; $p = .019$); and (4) washout component, 13.5% (AUC, 0.64; $p < .001$). The values to evaluate the association between high histologic grade and (1) tumor size was 1.85 cm (AUC, 0.66; $p < .001$); (2) angio-volume, 1.45 cm³ (AUC, 0.67; $p < .001$); (3) persistent component, 33.5% (AUC, 0.67; $p < .001$); (4) washout component, 15.5% (AUC, 0.73; $p < .001$); and (5) mean ADC value, 0.945×10^{-3} mm²/s (AUC, 0.61; $p = .008$).

Features associated with high Ki-67 and histologic grade

Univariate analysis demonstrated that high Ki-67 was associated with intratumoral high SI on T2WI, tumor size > 1.75 cm,

angio-volume >1.35 cm³, persistent component < 37.5%, and washout component > 13.5%. In multivariate analysis, washout component > 13.5% (odds ratio [OR] = 4.16; $p < .001$) and intratumoral high SI on T2WI (OR = 1.89; $p = .018$) were each associated with high Ki-67 (Table 2).

Univariate analysis was also conducted to determine if an association existed between high histologic grade and certain mpMRI features, and intratumoral high SI on T2WI, rim enhancement, axillary LN enlargement, tumor size > 1.85 cm, angio-volume > 1.45 cm³, persistent component < 33.5%, washout component > 15.5%, and mean ADC < 0.945×10^{-3} mm²/s were associated with high histologic grade. Multivariate analysis revealed that washout component > 15.5% (OR = 7.22; $p < .001$), rim enhancement (OR = 2.59; $p = .022$), and mean ADC < 0.945×10^{-3} mm²/s (OR = 2.47; $p = .015$) were each associated with high histologic grade (Table 3).

Diagnostic performance of prediction models

The diagnostic performances of prediction models using the eight machine learning algorithms are summarized in Table 4. Among eight methods using significant features from univariate analysis, two had the highest diagnostic performance for predicting high Ki-67 (AUC: 0.71) and k-nearest neighbor had the highest performance for

Table 4 AUCs of eight machine learning models using significant features acquired from univariate and multivariate logistic regression analysis for the prediction of high Ki-67 and histologic grade

	Prediction for high Ki-67		Prediction for high histologic grade	
	AUC	<i>p</i> value*	AUC	<i>p</i> value*
Machine learning models using significant features from univariate analysis				
Random forest	0.66 (0.57, 0.75)	0.093	0.69 (0.60, 0.78)	1.752e ⁻⁰⁷
Adaptive boosting	0.70 (0.62, 0.78)	0.544	0.80 (0.72, 0.89)	0.135
Decision tree	0.65 (0.57, 0.72)	0.032	0.57 (0.47, 0.67)	1.767e ⁻¹¹
k-nearest neighbor	0.69 (0.62, 0.75)	0.394	0.83 (0.76, 0.89)	Reference
Linear support vector machine	0.53 (0.43, 0.63)	0.003	0.72 (0.65, 0.80)	< 0.001
Naive Bayes classifier	0.69 (0.63, 0.75)	0.309	0.79 (0.73, 0.85)	0.073
Linear discriminant analysis	0.71 (0.63, 0.78)	0.404	0.80 (0.73, 0.88)	0.127
Logistic regression	0.71 (0.63, 0.78)	Reference	0.81 (0.72, 0.89)	0.140
Machine learning models using significant features from multivariate analysis				
Random forest	0.70 (0.61, 0.79)	0.746	0.74 (0.63, 0.85)	0.005
Adaptive boosting	0.70 (0.61, 0.79)	0.544	0.78 (0.70, 0.86)	0.441
Decision tree	0.70 (0.61, 0.79)	0.777	0.73 (0.63, 0.85)	0.004
k-nearest neighbor	0.68 (0.60, 0.77)	0.812	0.77 (0.68, 0.85)	0.035
Linear support vector machine	0.56 (0.44, 0.69)	< 0.001	0.50 (0.37, 0.62)	1.523e ⁻⁰⁵
Naive Bayes classifier	0.70 (0.61, 0.78)	0.294	0.79 (0.70, 0.88)	Reference
Linear discriminant analysis	0.70 (0.61, 0.78)	0.280	0.78 (0.69, 0.87)	0.462
Logistic regression	0.70 (0.61, 0.78)	Reference	0.78 (0.69, 0.87)	0.490

Numbers in parentheses are 95% confidence intervals

**P* value was acquired from comparison with the reference standard by using the DeLong method

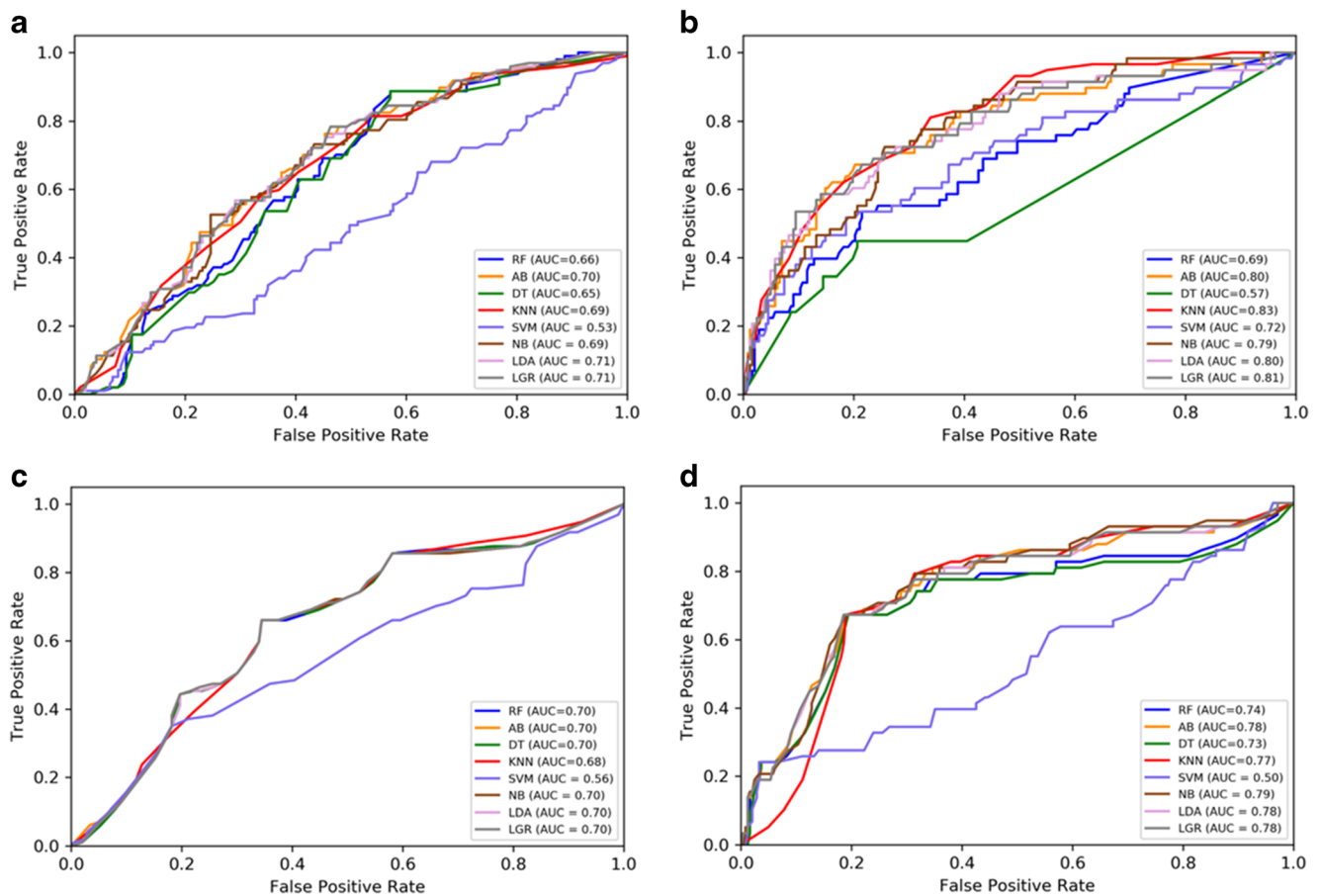


Fig. 3 Graphs show areas under receiver operating characteristic curve (AUCs) of eight machine learning algorithms using significant features acquired from univariate analysis for the prediction of higher Ki-67 (a) and higher histologic grade (b), and those using significant features acquired from multivariate analysis for the prediction of higher Ki-67

(c) and higher histologic grade (d). RF random forest, AB adaptive boosting, DT decision tree, kNN k-nearest neighbor, SVM linear support vector machine, NB naive Bayes, LDA linear discriminant analysis, LGR logistic regression

predicting high histologic grade (AUC: 0.83), whereas six had the highest performance for predicting high Ki-67 (AUC: 0.70) and naive Bayes classifier had the highest performance for predicting high histologic grade (AUC: 0.79) when using significant features from multivariate analysis (Fig. 3). Figure 4 is a box plot showing the diagnostic performance of eight machine learning methods.

Interobserver agreement

The overall agreement for ADC values between two readers was almost perfect ($r = 0.96$). For MRI morphologic imaging features, the overall agreement for FGT ($r = 0.62$), BPE ($r = 0.65$), mass internal enhancement ($r = 0.63$), associate NME ($r = 0.61$), intratumoral T2 high SI ($r = 0.72$), and axillary LN enlargement ($r = 0.78$) was substantial, while that for mass shape ($r = 0.53$), mass margin ($r = 0.57$), and peritumoral edema ($r = 0.56$) was moderate.

Discussion

Our study revealed that among various mpMRI features, CAD-measured washout component is key for predicting both higher Ki-67 and histologic grade in patients with stage I–II luminal breast cancer. A machine learning-based prediction model incorporating mpMRI features shows good diagnostic performance for predicting Ki-67 and histologic grade with AUCs of 70% and 79%, respectively.

Washout kinetics was the most important feature for this detailed prediction model. The results of our study are concordant with a prior study that indicated that the washout curve shape may predict a higher level of Ki-67 and histologic grade [23]. The impact of washout kinetics on poorer survival outcomes has previously been reported [24, 25]. One study found that a higher washout component was significantly associated with poorer disease-free survival [24], and another showed that patients with a smaller reduction in washout component after NAC had worse recurrence-free survival [25]. The most probable explanation for this is that hypervascular

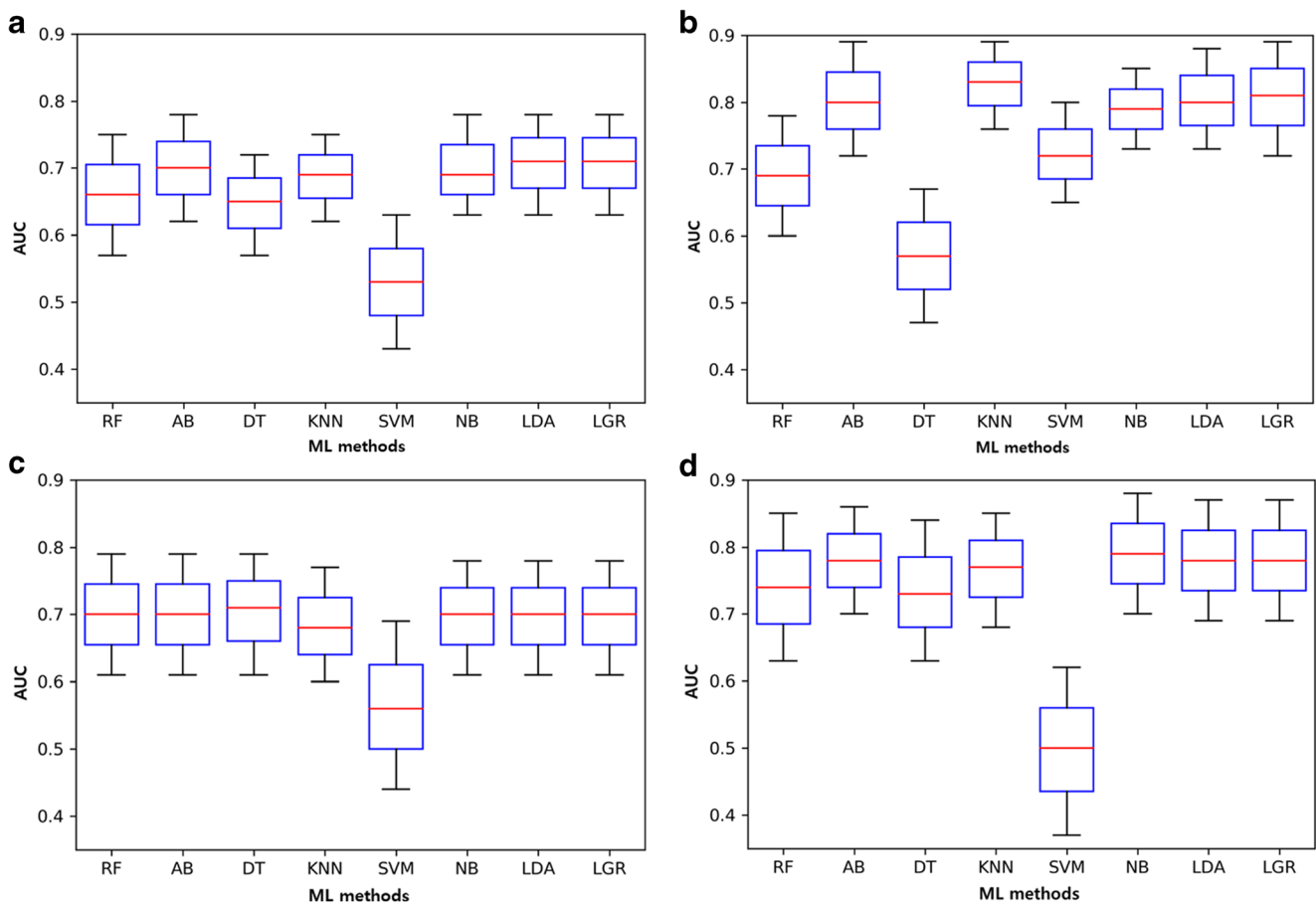


Fig. 4 Box and whisker plot illustrate diagnostic performances of eight machine learning (ML) algorithms using significant features acquired from univariate analysis for the prediction of higher Ki-67 (**a**) and higher histologic grade (**b**), and those using significant features acquired from multivariate analysis for the prediction of higher Ki-67

(**c**) and higher histologic grade (**d**). RF random forest, AB adaptive boosting, DT decision tree, kNN k-nearest neighbor, SVM linear support vector machine, NB naive Bayes, LDA linear discriminant analysis, LGR logistic regression

tumors may have a higher washout component indicating rapid perfusion because they have higher angiogenic markers, which could facilitate the spread of tumor cells into the body [26, 27]. Our study suggests that higher washout component could predict unfavorable prognosis due to its association with higher Ki-67 and histologic grade in luminal cancer [4].

The second-most influential mpMRI feature for predicting higher Ki-67 was the presence of intratumoral high SI on T2WI, not the ADC value. This result is similar to a recent study that found no correlation between mean ADC and Ki-67, while mean Ki-67 significantly increased with the amount of very high SI on T2WI, indicating tumor necrosis [28]. A possible explanation is that breast cancers with high proliferation (i.e., high Ki-67) may exceed the oxygen supply of their vascular system, resulting in prolonged hypoxia and subsequent necrosis with decreased cellularity [29]. Although controversy still remains regarding the correlation between ADC values and Ki-67 [14, 30–32], our

results support that intratumoral high SI on T2WI may be a more appropriate predictor for Ki-67 than the ADC value.

Contrary to Ki-67, lower ADC value can be used to predict higher histologic grade in luminal cancer. Previous studies also reported the inverse correlation between ADC value and histologic grade [31, 32]. However, another multicenter analysis of 870 breast cancer patients suggested that ADC cannot be used as a surrogate marker for Ki-67 and histologic grade [30]. Whereas the prior study attempted to distinguish grade 2 or 3 tumors from grade 1 tumors [30], we tried to distinguish grade 3 tumors from grade 1 or 2 tumors, because grade 3 luminal cancers are categorized one stage higher according to the eighth edition AJCC guidelines. From our results, we can assume that a higher histologic grade which means less tubule formation and higher mitotic activity could reflect higher cellularity, which results in lower ADC in luminal cancer.

Rim enhancement is a significant predictor of higher histologic grade [23, 32]. In luminal cancer subtype, the impact of rim enhancement on poorer distant metastasis-free survival has been reported [33]. In addition, a radiogenomic approach reported that a high calculated score indicating rim enhancement was correlated with early occurrence of metastasis and certain RNA phenotypes [34].

To date, machine learning algorithms have been applied to build a prediction model incorporating clinical or imaging data for further improving accuracy [35, 36]. In breast cancer patients, the integration of machine learning-based prediction models using mpMRI features has the potential to provide valuable information on predicting NAC treatment outcomes or risk of recurrence, and predicting which BI-RADS 3 and BI-RADS 4 lesions are benign [35–38]. Similar to the study of Eun et al [37], we used eight machine learning algorithms in our study to provide the best model to fit the input data and to predict correctly. Through comparison of the diagnostic performance of eight machine learning algorithms, we found that diagnostic performance using two or three significant features from multivariate analysis achieved the similar performance when using five or eight significant features from univariate analysis for predicting Ki-67 and histologic grade.

Our study had several limitations. First, this was a retrospective study from a single tertiary academic institution. Second, we excluded 25 lesions without enhancement on CAD and 24 lesions whose lesion visibility on DWI was too poor to acquire exact ADC measurements. Because the DWI acquisition resolution was low, it was difficult to measure ADC values in some subcentimeter lesions with poor lesion visibility. Therefore, selection bias may have been present in patient enrollment. Third, we did not analyze histopathological findings such as tumor necrosis and cellularity using the surgical specimens, which could have been important to support our findings. Fourth, the visual assessment of the presence of intratumoral high SI or peritumoral edema could be affected by postbiopsy hemorrhage because all MRI examinations were performed after imaging-guided biopsy. Fifth, we did not divide enrolled patients into test and validation sets. Instead, 10-fold cross-validation was used to train the eight machine learning algorithms to test the model's ability to predict correctly and to provide an insight into how the prediction model will generalize to an independent dataset.

In conclusion, higher values of CAD-measured washout component and a presence of high intratumoral SI on T2-weighted image were associated with high Ki-67, and higher values of CAD-measured washout component, rim enhancement, and lower apparent diffusion coefficient value were correlated with high histologic grade in luminal cancer subtypes. Our results provide an understanding

about the relationship between luminal cancer biology and mpMRI features. In addition, the integration of machine learning algorithms using mpMRI features can provide important predictive information to guide treatment decisions or determine the clinical prognosis for patients with stage I–II luminal cancers.

Funding This study has received funding from the Korea University Hospital Research Grant (K2023171).

Declarations

Guarantor The scientific guarantor of this publication is Kyu Ran Cho.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry Two of the authors have significant statistical expertise.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was obtained.

Methodology

- Retrospective
- Observational
- Performed at one institution

References

1. Gabriel NH, James LC, Carl JD et al (2017) Breast. In: Amin MB, American Joint Committee on Cancer (eds) AJCC cancer staging manual, 8th edn. Springer, New York, pp 589–628
2. Konecny G, Pauletti G, Pegram M et al (2003) Quantitative association between HER-2/neu and steroid hormone receptors in hormone receptor-positive primary breast cancer. *J Natl Cancer Inst* 95(2):142–153
3. Eiermann W, Rezai M, Kümmel S et al (2013) The 21-gene recurrence score assay impacts adjuvant therapy recommendations for ER-positive, node-negative and node-positive early breast cancer resulting in a risk-adapted change in chemotherapy use. *Ann Oncol* 24:618–624
4. Inwald EC, Koller M, Klinkhammer-Schalke M et al (2015) 4-IHC classification of breast cancer subtypes in a large cohort of a clinical cancer registry: use in clinical routine for therapeutic decisions and its effect on survival. *Breast Cancer Res Treat* 153:647–658
5. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ (2011) Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol* 22:1736–1747
6. Goldhirsch A, Winer EP, Coates AS et al (2013) Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol* 24:2206–2223

7. Weigelt B, Baehner FL, Reis-Filho JS (2010) The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. *J Pathol* 220:263–280
8. Onitilo AA, Engel JM, Greenlee RT, Mukesh BN (2009) Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. *Clin Med Res* 7:4–13
9. Hugh J, Hanson J, Cheang MC et al (2009) Breast cancer subtypes and response to docetaxel in node-positive breast cancer: use of an immunohistochemical definition in the BCIRG 001 trial. *J Clin Oncol* 27:1168–1176
10. Ehinger A, Malmström P, Bendahl PO et al (2017) Histological grade provides significant prognostic information in addition to breast cancer subtypes defined according to St Gallen 2013. *Acta Oncol* 56:68–74
11. Koh J, Kim MJ (2019) Introduction of a new staging system of breast cancer for radiologists: an emphasis on the prognostic stage. *Korean J Radiol* 20:69–82
12. Mori N, Ota H, Mugikura S et al (2015) Luminal-type breast cancer: correlation of apparent diffusion coefficients with the Ki-67 labeling index. *Radiology* 274:66–73
13. Surov A, Chang YW, Li L et al (2019) Apparent diffusion coefficient cannot predict molecular subtype and lymph node metastases in invasive breast cancer: a multicenter analysis. *BMC Cancer* 19:1043
14. Holli-Helenius K, Salminen A, Rinta-Kiikka I et al (2017) MRI texture analysis in differentiating luminal A and luminal B breast cancer molecular subtypes - a feasibility study. *BMC Med Imaging* 17:69
15. Grimm LJ, Zhang J, Mazurowski MA (2015) 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* 42:902–907
16. Pinker K, Helbich TH, Morris EA (2017) The potential of multiparametric MRI of the breast. *Br J Radiol* 90(1069):20160715
17. Morris EA, Comstock CE, Lee C et al (2013) ACR BI-RADS Magnetic resonance imaging. In: D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA (eds) ACR BI-RADS Atlas, breast imaging and reporting data system, 5th edn. Reston, American College of Radiology
18. Baltzer PA, Yang F, Dietzel M et al (2010) Sensitivity and specificity of unilateral edema on T2w-TSE sequences in MR-mammography considering 974 histologically verified lesions. *Breast J* 16:233–239
19. Uematsu T, Kasami M, Watanabe J (2014) Is evaluation of the presence of prepectoral edema on T2-weighted with fat-suppression 3 T breast MRI a simple and readily available noninvasive technique for estimation of prognosis in patients with breast cancer? *Breast Cancer* 21:684–692
20. Levman JE, Causer P, Warner E, Martel AL (2009) Effect of the enhancement threshold on the computer-aided detection of breast cancer using MRI. *Acad Radiol* 16:1064–1069
21. Hammond ME, Hayes DF, Dowsett M et al (2010) American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol* 28:2784–2795
22. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
23. Lee SH, Cho N, Kim SJ (2008) Correlation between high resolution dynamic MR features and prognostic factors in breast cancer. *Korean J Radiol* 9:10–18
24. Kim JJ, Kim JY, Kang HJ et al (2017) Computer-aided diagnosis-generated kinetic features of breast cancer at preoperative MR imaging: association with disease-free survival of patients with primary operable invasive breast cancer. *Radiology* 284:45–54
25. Yi A, Cho N, Im SA et al (2013) Survival outcomes of breast cancer patients who receive neoadjuvant chemotherapy: association with dynamic contrast-enhanced MR imaging with computer-aided evaluation. *Radiology* 268:662–672
26. Toi M, Inada K, Suzuki H, Tominaga T (1999) Tumor angiogenesis in breast cancer: its importance as a prognostic indicator and the association with vascular endothelial growth factor expression. *Breast Cancer Res Treat* 36:193–204
27. Uzzan B, Nicolas P, Cucherat M, Perret GY (2004) Microvessel density as a prognostic factor in women with breast cancer: a systematic review of the literature and meta-analysis. *Cancer Res* 64:2941–2955
28. Kim SY, Kim EK, Moon HJ et al (2018) Association among T2 signal intensity, necrosis, ADC and Ki-67 in estrogen receptor-positive and HER2-negative invasive ductal carcinoma. *Magn Reson Imaging* 54:176–182
29. Leek RD, Landers RJ, Harris AL, Lewis CE (1999) Necrosis correlates with high vascular density and focal macrophage infiltration in invasive carcinoma of the breast. *Br J Cancer* 79:991–995
30. Surov A, Clauser P, Chang YW et al (2018) Can diffusion-weighted imaging predict tumor grade and expression of Ki-67 in breast cancer? A multicenter analysis. *Breast Cancer Res* 20:58
31. Costantini M, Belli P, Rinaldi P et al (2010) Diffusion-weighted imaging in breast cancer: relationship between apparent diffusion coefficient and tumour aggressiveness. *Clin Radiol* 65:1005–1012
32. Kim KW, Kuzniak CM, Kim YJ, Seo JY, Jung HK, Lee MS (2018) Diagnostic usefulness of combination of diffusion-weighted imaging and T2WI, including apparent diffusion coefficient in breast lesions: assessment of histologic grade. *Acad Radiol* 25:643–652
33. Song SE, Shin SU, Moon HG, Ryu HS, Kim K, Moon WK (2017) MR imaging features associated with distant metastasis-free survival of patients with invasive breast cancer: a case-control study. *Breast Cancer Res Treat* 162:559–569
34. Yamamoto S, Maki DD, Korn RL, Kuo MD (2012) Radiogenomic analysis of breast cancer using MRI: a preliminary study to define the landscape. *AJR Am J Roentgenol* 199:654–663
35. Lo Gullo R, Eskreis-Winkler S, Morris EA, Pinker K (2020) Machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy. *Breast* 49:115–122
36. Reig B, Heacock L, Geras KJ, Moy L (2020) Machine learning in breast MRI. *J Magn Reson Imaging* 52:998–1018
37. Eun NL, Kang D, Son EJ et al (2020) Texture analysis with 3.0-T MRI for association of response to neoadjuvant chemotherapy in breast cancer. *Radiology* 294:31–41
38. Dalmış MU, Gubern-Mérida A, Vreemann S et al (2019) Artificial intelligence-based classification of breast lesions imaged with a multiparametric breast MRI protocol with ultrafast DCE-MRI, T2, and DWI. *Invest Radiol* 54:325–332

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.