AJR FOCUS ON: AJR JOURNAL CLUB

Frequency and Cancer Yield of BI-RADS Category 3 Lesions Detected at High-Risk Screening Breast MRI

Christine E. Edmonds[1]
Leslie R. Lamb
Sarah F. Mercaldo
Dorothy A. Sippo
Kristine S. Burk
Constance D. Lehman

OBJECTIVE. The purpose of this study was to evaluate the frequency and cancer yield of BI-RADS category 3 lesions in baseline versus nonbaseline (those with at least one prior) MRI screening examinations.

MATERIALS AND METHODS. Consecutive MRI screening examinations performed from 2011 through 2015 were reviewed. Pearson and Wilcoxon tests were used to examine differences in age, breast density, screening indication, background parenchymal enhancement, and cancer yield between baseline and nonbaseline MRI BI-RADS category 3 assessments. Multivariate logistic regression models based on generalized estimating equations were used to assess the odds of receiving a BI-RADS 3 assessment as a function of the variables.

RESULTS. Of 6672 MRI screening examinations of 3214 patients, 202 examinations (3%) were assessed BI-RADS category 3. Among baseline examinations, 8% (82/983) were assessed BI-RADS 3, compared with 2% (120/5689) of nonbaseline examinations ($p < 0.001$). Among the total BI-RADS 3 examinations, 6% (13/202) yielded malignancy of the lesion that had been assessed BI-RADS 3; 12 of 13 cancers were stage 0 or I at diagnosis. The cancer yield of BI-RADS 3 at baseline examinations was 2% (2/82), compared with 9% (11/120) for nonbaseline examinations ($p = 0.056$). Ten of 13 examinations were upgraded at or before 6-month follow-up MRI.

CONCLUSION. Baseline screening breast MRI examinations are associated with a significantly higher rate of BI-RADS category 3 assessments than are nonbaseline examinations. Most cancers diagnosed at follow-up of BI-RADS 3 lesions are in an early stage and are diagnosed at or before the 6-month follow-up examination. When used judiciously, short-interval follow-up MRI is an appropriate method for identifying early-stage breast cancer while avoiding unnecessary biopsies with benign findings.

Keywords: BI-RADS 3, breast cancer screening, breast MRI, probably benign

doi.org/10.2214/AJR.19.21778

Received May 28, 2019; accepted after revision July 26, 2019.

Based on a presentation at the ARRS 2019 Annual Meeting, Honolulu, HI.

Supported by a research grant (D. A. Sippo) and research support (C. D. Lehman) from GE Healthcare.

[1]All authors: Department of Radiology, Massachusetts General Hospital, Avon Comprehensive Breast Evaluation Center, Wang Ambulatory Care Bldg, Ste 240, 15 Parkman St, Boston, MA 02114. Address correspondence to C. E. Edmonds (cedmonds@mgh.harvard.edu).

This article is available for credit.

AJR 2020; 214:240–248

0361–803X/20/2142–240

Breast MRI is the most sensitive imaging modality for detecting breast cancer, facilitating early identification of malignancy that is mammographically and clinically occult [1–3]. The American Cancer Society, National Comprehensive Cancer Network, and American College of Radiology (ACR) all recommend MRI as an adjunct to screening mammography for women with a lifetime risk of 20% or greater as determined by hereditary risk models [4–6]. Since MRI screening guidelines for women at high risk of breast cancer were first issued in 2007, use of MRI for breast cancer screening has continued to increase, and screening continues to be the most common indication for breast MRI [7, 8].

The ACR BI-RADS atlas provides a standardized lexicon for describing breast findings and making diagnostic assessments and recommendations about mammography, sonography, and MRI. BI-RADS category 3 was initially established for mammographic lesions that were thought to be probably benign and to carry an estimated cancer risk of 2% or less. The purpose was to limit the number of breast biopsies performed for findings that cannot be definitively characterized as benign on mammograms but ultimately have benign pathologic results [9]. The few malignancies diagnosed during follow-up of probably benign lesions should be small, early stage, and predominantly node negative [9]. The recommended management of probably benign mammographic lesions is short-interval follow-up imaging, typically performed 6, 12, and 24 months after the initial imaging [9].

The appropriate utilization and outcomes of the mammographic BI-RADS category 3

assessment have been thoroughly validated in multiple research studies [9–13]. By contrast, for MRI BI-RADS 3 lesions, the expectation of a malignancy rate of 2% or less and the follow-up imaging recommendations are borrowed from mammography. There is a relative paucity of data with which to validate and guide appropriate use of BI-RADS category 3 in MRI. Although the U.S. Food and Drug Administration does not currently mandate auditing of breast MRI practice as it does mammography, the ACR mandates an auditing program for breast MRI accreditation, and the most recent BI-RADS atlas [14] has been updated to include breast MRI screening benchmarks.

The most recent edition of the BI-RADS manual [14] suggests a goal for frequency of BI-RADS 3 assessments of less than 10%. It also notes that as a breast imaging program gains experience and matures over time, this rate should decrease to the approximately 1–2% rate achieved for mammography. However, research to date shows great practice variability in the use of BI-RADS 3 in MRI, with frequencies of use ranging from 6% to 24% and malignancy rates from 0% to 10% [15–27]. This variability in utilization and outcomes is likely related to small study sample sizes, variability in MRI experience, and heterogeneous subject populations that often represent a combination of screening and diagnostic MRI indications. Multiple studies to date have shown differences in various performance measures, including rates of abnormal interpretation and cancer detection, for screening versus diagnostic breast MRI, supporting the need for separate data to guide screening versus diagnostic MRI practice [28–30].

Given the highly variable data on BI-RADS 3 utilization and outcomes, further research is warranted to assess and optimize application of BI-RADS category 3 in a large high-risk screening population. In this study, we aimed to evaluate the frequency of BI-RADS 3 assessment in a pure high-risk screening population, compare the frequency of BI-RADS 3 assessment in baseline examinations with the frequency among nonbaseline screening examinations, and compare the cancer yield of BI-RADS 3 lesions in patients undergoing baseline screening compared with those undergoing nonbaseline (defined as screening examinations with at least one prior breast MRI examination) screening examinations.

Materials and Methods

Study Population

Our institutional review board approved this retrospective case review study with waiver of the requirement for informed consent. In compliance with HIPAA, we reviewed all consecutive screening breast MRI examinations performed from January 1, 2011, through December 31, 2015 (n = 6729). Examinations performed for atypical screening indications, such as mammographically dense breasts, were excluded (n = 57). This yielded a final screening breast MRI cohort of 6672 examinations of 3214 women.

Data Sources

The following data were obtained from our breast imaging information system (MagView Mammography, MagView) and from review of the electronic medical record: patient demographics, primary screening indication, mammographic screening density, qualitative background parenchymal enhancement (BPE) prospectively reported at clinical MRI interpretation, BI-RADS assessment, availability of prior comparison MRI examinations, year of examination, and pathologic result.

For patients with multiple indications for high-risk screening, the following hierarchy, starting with the highest risk, was used to define the primary indication based on the associated level of risk: *BRCA* mutation carrier (BRCA), history of thoracic radiation, personal history of breast cancer, personal history of high-risk lesion, and family history of breast cancer. Mammographic density was recorded from the most contemporaneous mammogram within 24 months of the MRI examination, because the amount of fibroglandular tissue was not consistently stated in the MRI reports. BPE was reported according to the four categories of minimal, mild, moderate, and marked as defined in the ACR 2013 BI-RADS atlas [14].

MRI Technique

MR images were obtained in the axial plane with an acquired slice thickness of 3 mm or less at 1.5 T or 3 T with a dedicated breast coil and the patient prone. The protocol followed the recommendations of the ACR for an unenhanced non–fat-suppressed T1-weighted sequence and an unenhanced fat-suppressed T2-weighted sequence. In addition, an unenhanced fat-suppressed gradient-echo T1-weighted sequence was performed and followed by two to four dynamic contrast-enhanced T1-weighted gradient-echo series with fat suppression after IV administration of a gadolinium-based contrast agent. Postprocessing included sagittal and coronal reconstructions, subtracted contrast-enhanced images, and maximum-intensity-projection images. Examinations were interpreted by fellowship-trained or equivalent (defined as a minimum of 5 years' clinical experience interpreting breast imaging) breast imagers using the terminology of the BI-RADS atlas.

Performance Metrics

BI-RADS assessments were collected from the MRI reports. At our institution, breast MRI examinations are assigned a single overall BI-RADS assessment for both breasts. Examinations were excluded from the screening cohort if they were performed for follow-up of a previous BI-RADS 3 assessment (defined as performed within 14 months of a previous MRI examination with a BI-RADS 3 assessment). All biopsy and surgical pathologic results obtained within 2 years of a BI-RADS 3 assessment were reviewed. In addition, the cases of all patients with BI-RADS 3 assessments without follow-up MRI reports in our breast imaging database system were reviewed in the electronic medical record to determine adherence to follow-up breast imaging recommendations and to determine the number of patients with BI-RADS 3 assessments lost to follow-up. BI-RADS 3 assessments without malignant pathologic results within 2 years after the MRI date were considered benign.

For all BI-RADS 3 assessments with resulting breast malignancy within 2 years of MRI, imaging reports were reviewed to determine whether the BI-RADS 3 lesion was the site of the resulting malignancy. If it was unclear from the imaging reports alone whether the resulting malignancy site corresponded to the lesion that received a BI-RADS 3 assessment, the MR images were also reviewed. BI-RADS 3 cases with resulting malignancy at the site of the BI-RADS 3 lesion within 2 years after MRI were considered positive for malignancy. Malignancy was defined as invasive carcinoma or ductal carcinoma in situ within the breast. For patients with multiple lesions, the lesions with the highest-severity pathologic result was recorded (for example, invasive carcinoma ranked higher than ductal carcinoma in situ). For these positive cases, details were collected from the electronic medical record regarding the MRI lesion descriptor, the pathologic diagnosis and stage, and the time to diagnosis and mode of diagnosis. The time to diagnosis was considered the period in months from the BI-RADS 3 MRI assessment to follow-up imaging (MRI, ultrasound, or mammography) that initiated biopsy or the time to biopsy (in one case in which no further imaging preceded the biopsy).

Statistical Analyses

The Wilcoxon test (for continuous variables) and the Pearson chi-square test (for categoric vari-

ables) were used to evaluate the distribution of age, breast density, baseline status (baseline examination vs nonbaseline examination), screening indication, BPE, and final cancer diagnosis in the cohort of examinations assessed BI-RADS 3 compared with all other BI-RADS assessments combined. For the analysis, moderate and marked BPE were combined into a single moderate-marked group, for a comparison of minimal versus mild versus moderate-marked. Examinations performed for an indication of BRCA or thoracic radiation were combined into one screening indication, given the small numbers of examinations performed for this indication and the similar high degree of breast cancer risk associated with each indication [31].

Multivariable logistic regression models were estimated by means of generalized estimating equations with an independent correlation structure and robust sandwich standard errors to account for multiple examinations of one patient. For this analysis, BPE was dichotomized into minimal-mild and moderate-marked groups. Mammographic breast density was also dichotomized into dense (heterogeneously dense or extremely dense) and not dense (almost entirely fatty or scattered fibroglandular densities). The model is used to examine the odds of a BI-RADS 3 assessment as a function of baseline MRI status (baseline vs nonbaseline), age, breast density (dense vs not dense), screening indication (family history of breast cancer, personal history of high-risk lesion, personal history of breast cancer vs BRCA or thoracic radiation), BPE (moderate-marked vs minimal-mild), and year (2012, 2013, 2014, and 2015 vs 2011). Only cases with all relevant data were included, resulting in 5957 examinations used for model estimation. Adjusted odds ratio (OR), 95% CI, and the corresponding Wald p were calculated. Type I error of 5% was used for all CIs and hypothesis tests. All data were analyzed with statistical software (R version 3.5.1, R, Foundation for Statistical Computing). The packages rms and geepack were used in this analysis.

Results

Among the 6672 examinations of 3214 patients, 202 (3%) examinations were assessed BI-RADS category 3. Of the 202 BI-RADS 3 assessments, 82 (41%) were for baseline examinations and 120 (59%) were for nonbaseline examinations. Among the baseline screening examinations, 8% (82/983) were assessed BI-RADS 3, compared with 2% (120/5689) of nonbaseline examinations ($p < 0.001$). Nine of the 202 (4%) patients with BI-RADS 3 assessments did not adhere to follow-up imaging recommendations and never underwent follow-up breast MRI or had pathologic proof of benignity of the BI-RADS 3 lesion. However, eight of these nine continued to undergo mammography or clinical follow-up at our institution. Only one of the nine was entirely lost to follow-up at our institution after the BI-RADS 3 MRI examination. The other 193 BI-RADS 3 cases (96%) were deemed benign at follow-up imaging or pathologic analysis (biopsy or prophylactic mastectomy).

Table 1 summarizes the characteristics of the screening MRI examinations, comparing examinations with BI-RADS 3 assessments against the cohort of examinations with non–BI-RADS 3 assessments (BI-RADS 0, 1, 2, 4, or 5). Patients with a BI-RADS 3 assessment were younger (median, 47.0 years; interquartile range [IQR], 41.0–54.0 years) than those with a non–BI-RADS 3 assessment (median, 52.0 years; IQR, 46.0–59.0 years) ($p < 0.001$). There was no evidence of a difference in breast density between the BI-RADS 3 cohort and the non–BI-RADS 3 cohort. The most common screening indication among both the BI-RADS 3 cohort and the non–BI-RADS 3 cohort was personal his-

TABLE 1: Characteristics of 6672 Screening Breast MRI Examinations According to BI-RADS Assessment

Characteristic	No.	BI-RADS Category			p
		All (n = 6672)	0–2, 4, 5 (n = 6470)	3 (n = 202)	
Year of examination	6672				0.003
2011		1465 (22)	1402 (22)	63 (31)	
2012		1359 (20)	1320 (20)	39 (19)	
2013		1151 (17)	1130 (17)	21 (10)	
2014		1195 (18)	1166 (18)	29 (14)	
2015		1502 (23)	1452 (22)	50 (25)	
Age (y)	6672				< 0.001
Lower quartile		46	46	41	
Median		52	52	47	
Upper quartile		58	59	54	
Mammographic breast density[a]	6436				0.226
Fatty		153 (2)	148 (2)	5 (3)	
Scattered		2000 (31)	1952 (31)	48 (25)	
Heterogeneously dense		3475 (54)	3364 (54)	111 (57)	
Extremely dense		808 (13)	778 (12)	30 (15)	
Primary indication	6672				0.009
BRCA, thoracic radiation		786 (12)	760 (12)	26 (13)	
Family history		1812 (27)	1746 (27)	66 (33)	
High-risk lesion		694 (10)	664 (10)	30 (15)	
Personal history		3380 (51)	3300 (51)	80 (40)	
Background parenchymal enhancement[b]	6170				< 0.001
Minimal		2301 (37)	2256 (38)	45 (24)	
Mild		2810 (46)	2723 (46)	87 (46)	
Moderate-marked		1059 (17)	1001 (17)	58 (31)	
Baseline MRI status	6672				
Nonbaseline		5689 (85)	5569 (86)	120 (59)	< 0.001
Baseline		983 (15)	901 (14)	82 (41)	

Note—Except for age, data are number of examinations with percentage in parentheses. BRCA = *BRCA* mutation carrier.

[a]For breast density, 236 examinations were not available.

[b]For background parenchymal enhancement, 502 examinations were not available.

TABLE 2: Characteristics of 202 BI-RADS Category 3 Examinations According to Baseline Status

Characteristic	No.	Baseline (*n* = 82)	Nonbaseline (*n* = 120)	*p*
Year of examination	202			0.298
2011		21 (26)	42 (35)	
2012		13 (16)	26 (22)	
2013		9 (11)	12 (10)	
2014		15 (18)	14 (12)	
2015		24 (29)	26 (22)	
Age (y)	202			0.005
Lower quartile		39.0	43.0	
Median		45.0	48.5	
Upper quartile		51.0	56.0	
Mammographic breast density[a]	194			0.245
Fatty		3 (4)	2 (2)	
Scattered		19 (25)	29 (25)	
Heterogeneously dense		39 (51)	72 (62)	
Extremely dense		16 (21)	14 (12)	
Primary indication	202			< 0.001
BRCA, thoracic radiation		14 (17)	12 (10)	
Family history		41 (50)	25 (21)	
High-risk lesion		15 (18)	15 (12)	
Personal history		12 (15)	68 (57)	
Background parenchymal enhancement[b]	190			0.052
Minimal		17 (21)	28 (25)	
Mild		31 (39)	56 (51)	
Moderate-marked		32 (40)	26 (24)	
Cancer diagnosis	202			0.056
Cancer		2 (2)	11 (9)	
No cancer		80 (98)	109 (91)	

Note—Data are number of examinations with percentage in parentheses. BRCA = *BRCA* mutation carrier.
[a]For breast density, eight examinations were not available.
[b]For background parenchymal enhancement, 12 examinations were not available.

tory of breast cancer. We also found that 31% (58/190) of BI-RADS 3 examinations showed moderate or marked BPE compared with 17% (1001/5980) of non–BI-RADS 3 examinations ($p < 0.001$). Among the total BI-RADS 3 assessments, 41% (82/202) were of baseline examinations. Only 14% (901/6470) of the non–BI-RADS 3 assessments were of baseline examinations ($p < 0.001$).

Table 2 compares the characteristics of the 82 BI-RADS 3 assessments of baseline MRI examinations with the 120 BI-RADS 3 assessments of nonbaseline examinations. Patients with a BI-RADS 3 assessment of a baseline examination were younger (median, 45.0 years; IQR, 39.0–51.0 years) than the patients with a BI-RADS 3 assessment of a nonbaseline examination (median, 48.5 years; IQR, 43.0–56.0 years) ($p = 0.005$). There was no evidence of a difference in breast density between baseline and nonbaseline BI-RADS 3 examinations. We found that 40% (32/80) of baseline BI-RADS 3 examinations showed moderate or marked BPE compared with 24% (26/110) of nonbaseline BI-RADS 3 examinations ($p = 0.052$). Of the total BI-RADS 3 assessments of screening MRI, 6% (13/202) yielded malignancy; 2% (2/82) of baseline BI-RADS 3 examinations yielded malignancy, compared with 9% (11/120) of nonbaseline BI-RADS 3 examinations ($p = 0.056$).

The multivariable logistic regression generalized estimating equation analysis (Table 3) showed that baseline screening MRI examinations are significantly more likely to receive a BI-RADS 3 assessment than are nonbaseline examinations (adjusted OR, 4.17; 95% CI, 2.99–5.81; $p < 0.001$) after adjustment for age, BPE, breast density, screening indication, and year. In addition, a BPE assessment of moderate-marked was significantly associated with increased odds of a BI-RADS 3 assessment at screening MRI compared with minimal-mild BPE (adjusted OR, 1.66; 95% CI, 1.17–2.35; $p = 0.004$) after adjustment for the other variables. The odds of receiving a BI-RADS 3 assessment in each year after 2011 (2012–2015) were significantly lower compared with 2011.

Table 4 characterizes the 13 BI-RADS 3 MRI examinations that yielded cancer and the details on the cancers. Of the 13 MRI examinations, the screening indication was personal history of breast cancer for nine examinations (69%), personal history of high-risk lesion for two (15%), BRCA for one (8%), and family history of breast cancer for one (8%). Eight of 13 (62%) findings were characterized as a focus or foci (Fig. 1), three (23%) as nonmass enhancement, and two (15%) as masses. Seven of the malignant diagnoses were made at 6-month follow-up MRI examinations (all were performed 6–7 months from the date of the BI-RADS 3 MRI assessment), two were made at approximately 12-month follow-up MRI (11 and 13 months), and one was made at screening MRI performed 24 months after the BI-RADS 3 assessment (this case was downgraded to BI-RADS category 2 at the first follow-up MRI examination but upgraded to BI-RADS category 4 at subsequent screening MRI). One of the other three patients with malignancies underwent MRI-guided biopsy of the BI-RADS 3 lesion at an outside hospital 3 months after the BI-RADS 3 assessment, despite the recommendation for follow-up MRI. The second underwent diagnostic mammography and ultrasound because of an area of clinical concern that corresponded to the same location as the BI-RADS 3 lesion 2 months after the BI-RADS 3 assessment, and the assessment was upgraded to BI-RADS 4. The third patient underwent diagnostic mammography and ultrasound of an area of clinical concern that corresponded to the location of the BI-RADS 3 lesion 3 months after the BI-RADS 3 assessment, and the assessment was upgraded to BI-RADS 4.

TABLE 3: Adjusted Odds Ratios and 95% CIs for All Model Covariates (*n* = 5957)

Variable	Odds Ratio	*p*
Age	0.98 (0.97–1.00)	0.027
Background parenchymal enhancement		
Minimal-mild	Reference	
Moderate-marked	1.66 (1.17–2.35)	0.004
Breast density		
Fatty, scattered	Reference	
Heterogeneously, extremely dense	1.21 (0.86–1.71)	0.278
Screening indication		
BRCA, thoracic radiation	Reference	
Family history	0.88 (0.52–1.51)	0.651
High-risk lesion	1.41 (0.77–2.60)	0.271
Personal history	1.17 (0.70–1.94)	0.548
Examination year		
2011	Reference	
2012	0.59 (0.38–0.92)	0.021
2013	0.38 (0.22–0.65)	< 0.001
2014	0.49 (0.30–0.79)	0.003
2015	0.55 (0.36–0.85)	0.007
Baseline MRI status		
Nonbaseline	Reference	
Baseline	4.17 (2.99–5.81)	< 0.001

Note—Values in parentheses are 95% CI. The model is used to examine the odds of being assessed BI-RADS category 3 as a function of baseline MRI (baseline vs not baseline), age, breast density (dense vs not dense), screening indication (family history, high-risk lesion, personal history vs *BRCA* mutation carrier [BRCA], thoracic radiation), background parenchymal enhancement (moderate-marked vs minimal-mild), and year (2012, 2013, 2014, 2015 vs 2011).

Discussion

Although there are extensive data supporting both the use of a BI-RADS category 3 assessment for specific mammographic and sonographic findings and the recommendations for imaging follow-up, data guiding the use and follow-up of a BI-RADS 3 assessment at breast MRI are limited [14]. According to the ACR recommendations, a finding assessed BI-RADS 3 at MRI should have 2% or less chance of malignancy but greater than essentially 0% chance of malignancy on the basis of similar recommendations for mammography and sonography [14]. In addition, the follow-up imaging recommendations for BI-RADS 3 assessments at MRI are borrowed from mammography recommendations: MRI follow-up 6, 12, 24, and optionally 36 months after the initial imaging is typically recommended to establish stability [14].

In the current study, we sought to measure the rate of BI-RADS 3 assessments in a large database of MRI examinations performed purely for high-risk screening and to determine whether utilization of BI-RADS 3 varies between baseline and nonbaseline screening examinations. We also sought to determine whether the cancer yield of BI-RADS 3 lesions at screening MRI differs between baseline and nonbaseline MRI examinations. The aim was to use the results to better inform application of the BI-RADS category 3 assessment in high-risk screening MRI.

TABLE 4: Characteristics of 13 Cases of BI-RADS Category 3 Assessments Yielding Malignancy

Baseline Examination Performed	MRI Descriptor	Screening Indication	Time to Diagnosis (mo)	Mode of Diagnosis	Diagnosis	Stage at Diagnosis
Yes	Oval mass	*BRCA* mutation	2	Diagnostic mammography and ultrasound	IDC	I
Yes	Mass	Personal history	6	Follow-up MRI	IDC	III
No	Focus	Personal history	6	Follow-up MRI	IDC	I
No	Focus	Personal history	6	Follow-up MRI	IDC	I
No	Focus	Personal history	11	Follow-up MRI	IDC	I
No	Focus	Personal history	3	MRI biopsy	ILC	I
No	Focus	Personal history	3	Diagnostic mammography and ultrasound	IDC	I
No	Focus	Personal history	6	Follow-up MRI	IDC, ILC	I
No	Focus	High-risk lesion	6	Follow-up MRI	DCIS	0
No	Multiple foci	Personal history	13	Follow-up MRI	DCIS	0
No	Regional NME	Family history	7	Follow-up MRI	IDC	I
No	Linear NME	Personal history	6	Follow-up MRI	DCIS with microinvasion	I
No	NME	High-risk lesion	24	Screening MRI	DCIS	0

Note—IDC = invasive ductal carcinoma, ILC = invasive lobular carcinoma, DCIS = ductal carcinoma in situ, NME = nonmass enhancement.

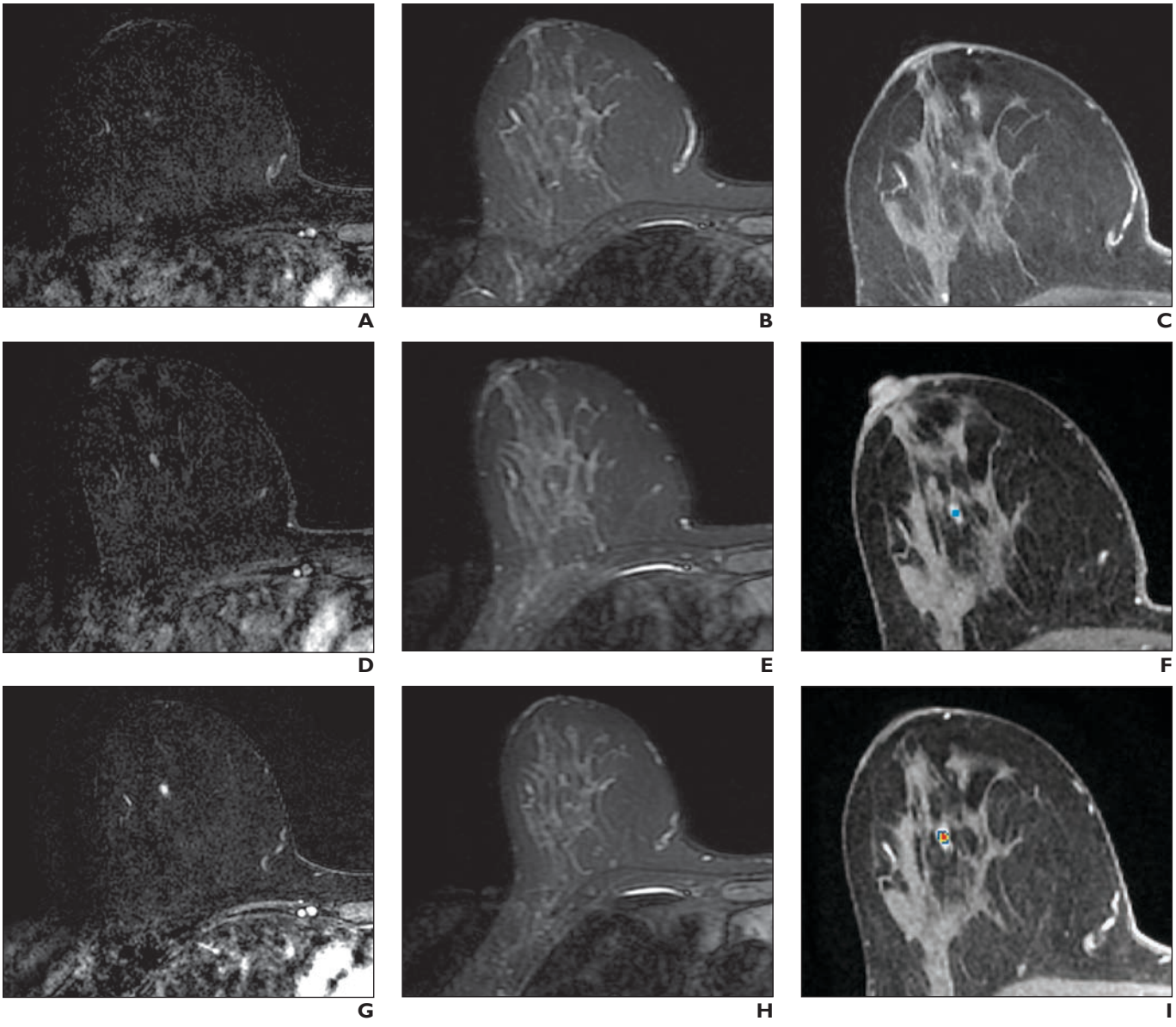

Fig. 1—50-year-old woman undergoing high-risk screening MRI with history of left breast atypical ductal hyperplasia after excision and family history of breast cancer. Only right breast is shown.

A–C, Screening axial contrast-enhanced fat-suppressed T1-weighted subtraction image (**A**), axial unenhanced fat-suppressed T2-weighted image (**B**), and axial contrast-enhanced fat-suppressed T1-weighted image with kinetic color overlay map (**C**) show no suspicious findings. Examination was assessed BI-RADS category 2.

D–F, Nonbaseline screening axial contrast-enhanced fat-suppressed T1-weighted subtraction image (**D**), axial unenhanced fat-suppressed T2-weighted image (**E**), and axial contrast-enhanced fat-suppressed T1-weighted image with kinetic color overlay map (**F**) obtained approximately 1 year after **A–C**. Contrast-enhanced images show focus of enhancement (**D** and **F**) in central right breast without T2 correlate (**E**) and with persistent kinetics (*blue*) (**F**). Focus was assessed BI-RADS 3 with recommendation for 6-month follow-up breast MRI.

G–I, Axial contrast-enhanced fat-suppressed T1-weighted subtraction image (**G**), axial unenhanced fat-suppressed T2-weighted image (**H**), and axial contrast-enhanced fat-suppressed T1-weighted image with kinetic color overlay map (**I**) obtained 6 months after **D–F** show increased conspicuity of right breast lesion and washout kinetics (red represents washout; yellow, plateau; blue, persistent delayed phase kinetics). Lesion was upgraded to BI-RADS 4 at this examination, and MRI-guided biopsy was recommended. Both MRI biopsy and excisional biopsy yielded grade 2 ductal carcinoma in situ.

In our high-risk screening cohort, 3% of examinations were assessed BI-RADS 3. This percentage is lower than that found in previous studies, in which the frequency of BI-RADS 3 assessments ranged from 6% to 24% [15–26]. Most of the prior studies were conducted with more heterogeneous cohorts that included both screening and diagnostic breast MRI examinations. In addition, both breast MRI technique and reader experience have notably improved since earlier studies. Relevant studies performed from 2014 to 2017 [15, 18, 22, 24, 26] showed BI-RADS 3 frequencies of 5–8.5%. Our frequency of 3% approaches the goal of 1–2% recommended in the ACR 2013 BI-RADS atlas [14] for a mature and experienced MRI screening program.

Prior studies [15–26] have shown a wide range (0% to 10%) of malignancy rates among MRI BI-RADS 3 lesions. Again, this variation likely reflects heterogeneity among

the study cohorts. Many of the previous studies [15, 16, 18, 22, 23, 32] included a combined population of both screening and diagnostic examinations. Weinstein et al. [25] examined the frequency of malignancy of BI-RADS 3 lesions in the prospective ACR American College of Radiology Imaging Network 6667 cohort and found a very low BI-RADS 3 malignancy rate of 0.9%; however, this cohort included only subjects with a recent diagnosis of breast cancer and thus was not representative of a true screening population. In another study of the malignancy rate among 108 BI-RADS 3 lesions, Spick et al. [24] also found a very low malignancy rate of 0.9%. That study, however, also did not reflect a true high-risk screening population, because it excluded all subjects with a personal history of breast cancer or who were at high risk of breast cancer.

To our knowledge, the current study of MRI BI-RADS 3 utilization and outcomes in a high-risk screening population is the largest to date. We found an overall BI-RADS 3 cancer yield of 6%. Although this rate is somewhat higher than that found in previous studies, our study population is unique in that it is a pure high-risk population based on well-established high-risk screening indications. The 6% cancer yield is also notably higher than the ACR recommended target ceiling rate of 2%, borrowed from mammographic and sonographic data. We postulate that the 2% or less BI-RADS malignancy rate for mammography may not represent the ideal malignancy rate for high-risk screening MRI. The population undergoing screening MRI is at substantially higher risk of breast cancer than is the general population undergoing screening mammography; therefore, a higher percentage of BI-RADS 3 lesions may be expected to ultimately yield malignancy.

As with mammography, the primary goal of the BI-RADS category 3 in MRI is to avoid unnecessary breast biopsies, that is, to decrease the number of biopsies performed for findings that cannot definitely be characterized as benign but ultimately yield benign pathologic results. In our study, 189 of the 202 BI-RADS 3 assessments did not yield cancer, and unnecessary biopsy was therefore avoided. Thus, 189 of the 6672 patients undergoing screening, or 2.8% of the screening population, avoided an unnecessary biopsy.

Our data suggest that we may safely observe BI-RADS 3 MRI lesions and accept a higher malignancy rate without risking diagnosing malignancy at more advanced stages. Twelve of the 13 BI-RADS 3 cancers identified in our study were early-stage node-negative disease at diagnosis: three cases of stage 0 and nine of stage I disease. The thirteenth cancer was diagnosed at 6-month follow-up MRI, and the imaging findings were suggestive of stage III disease at diagnosis. In retrospect, however, this case did not meet criteria for a BI-RADS 3 assessment and was likely already at an advanced disease stage at initial screening MRI. In addition to a breast mass that ultimately yielded malignancy, an enlarged internal mammary node ipsilateral to the mass was present and was suspicious for nodal metastasis. Taken together, our data raise the question whether breast imagers may safely accept a malignancy rate higher than the 2% cutoff borrowed from mammography to reduce false-positive biopsy recommendations while maintaining an early-stage breast cancer diagnosis. Further investigation in this arena is warranted.

Our investigation was powered to assess differences in utilization of BI-RADS 3 in baseline versus nonbaseline examinations. The odds of a BI-RADS 3 assessment in a baseline examination were approximately four times that of a BI-RADS 3 assessment in a nonbaseline examination (Table 4). To our knowledge, this difference in utilization has not previously been examined. This difference makes intuitive sense given the important role that stability of a finding plays in suggesting benignity at MRI. Furthermore, the results of this study suggest that there may be a difference in cancer yield between BI-RADS 3 assessments in baseline and nonbaseline examinations, nonbaseline BI-RADS 3 assessments yielding notably higher rates of malignancy (9%) than baseline BI-RADS 3 assessments (2%). Although this difference did not quite show statistical significance ($p = 0.056$), the results suggest that a larger sample size may definitively show higher malignancy rates in nonbaseline compared with baseline BI-RADS 3 assessments. On the basis of these data, breast imagers should exercise greater caution in using BI-RADS 3 for nonbaseline examinations. Although use of BI-RADS 3 for baseline examinations yields an acceptably low malignancy rate, further investigation is warranted to determine whether there is an appropriate role for safe use of BI-RADS 3 for nonbaseline examinations, perhaps based on specific lesion features, or whether BI-RADS 3 should be used exclusively in baseline examinations.

Finally, our investigation offers insight into the optimal timing of follow-up of MRI BI-RADS 3 assessments. Given the high cost of breast MRI, the variable insurance coverage of more than one MRI examination per year, and the historically low concordance rates between MRI BI-RADS 3 assessments and short-interval follow-up imaging recommendations [33], the utility of 6-month follow-up MRI, as opposed to follow-up at 1 year, is an important clinical question. Ten of the 13 malignant BI-RADS 3 lesions were identified at or before the 6-month follow-up MRI study, suggesting that 6-month follow-up MRI is critical to early diagnosis.

Although our investigation had the advantages of a large and homogeneous high-risk screening population, there were limitations. Our data were not linked to a tumor registry. Although we reviewed the medical records of all patients with BI-RADS 3 assessments who did not have follow-up imaging in our system, it is possible that we did not identify all of the malignant BI-RADS 3 lesions. In addition, our experience at a tertiary-care academic institution may not translate appropriately to community breast screening practices. Finally, over the 5 years of the study period, the MRI protocol at our institution evolved as the technology improved, we added a 3-T MRI system, and our experience grew. Therefore, some details of the protocol, such as the number of contrast-enhanced sequences and the exact timing of the sequences, evolved over the years of the study. The heterogeneity of these protocols was a possible limitation of this research.

Conclusion

Our study results confirm that it is feasible for a mature breast MRI screening program to use BI-RADS category 3 at a low overall rate of 3%, approaching the 1–2% recommendation for mammography. Our data also establish that baseline screening MRI examinations are associated with a significantly higher rate of BI-RADS 3 assessments than nonbaseline examinations and suggest that BI-RADS 3 lesions in nonbaseline examinations have a higher cancer yield than do those in baseline examinations. Given that most cancers diagnosed during follow-up of BI-RADS 3 lesions are early stage, there is strong evidence that judicious use of short-interval follow-up MRI is an appropriate method for identifying early-stage breast cancer while avoiding unnecessary biopsies of benign breast lesions. However, greater

prudence is warranted in the assessment of BI-RADS category 3 in nonbaseline examinations, and further investigation is warranted to better guide appropriate utilization.

References

1. Lehman CD, Isaacs C, Schnall MD, et al. Cancer yield of mammography, MR, and US in high-risk women: prospective multi-institution breast cancer screening study. *Radiology* 2007; 244:381–388
2. DeMartini W, Lehman C. A review of current evidence-based clinical applications for breast magnetic resonance imaging. *Top Magn Reson Imaging* 2008; 19:143–150
3. Warner E, Messersmith H, Causer P, Eisen A, Shumak R, Plewes D. Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. *Ann Intern Med* 2008; 148:671–679
4. National Comprehensive Cancer Network. Breast caner. In: *NCCN clinical practice guidelines in oncology (NCCN guidelines)*. Vol. 3. Fort Washington, PA: National Comprehensive Cancer Network, 2018
5. Saslow D, Boetes C, Burke W, et al. American Cancer Society Breast Cancer Advisory Group. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin* 2007; 57:75–89
6. Monticciolo DL, Newell MS, Moy L, Niell B, Monsees B, Sickles EA. Breast cancer screening in women at higher-than-average risk: recommendations from the ACR. *J Am Coll Radiol* 2018; 15(3 pt A):408–414
7. Wernli KJ, DeMartini WB, Ichikawa L, et al.; Breast Cancer Surveillance Consortium. Patterns of breast magnetic resonance imaging use in community practice. *JAMA Intern Med* 2014; 174:125–132
8. Stout NK, Nekhlyudov L, Li L, et al. Rapid increase in breast magnetic resonance imaging use: trends from 2000 to 2011. *JAMA Intern Med* 2014; 174:114–121
9. Sickles EA. Periodic mammographic follow-up of probably benign lesions: results in 3,184 consecutive cases. *Radiology* 1991; 179:463–468
10. Helvie MA, Pennes DR, Rebner M, Adler DD. Mammographic follow-up of low-suspicion lesions: compliance rate and diagnostic yield. *Radiology* 1991; 178:155–158
11. Varas X, Leborgne F, Leborgne JH. Nonpalpable, probably benign lesions: role of follow-up mammography. *Radiology* 1992; 184:409–414
12. Varas X, Leborgne JH, Leborgne F, Mezzera J, Jaumandreu S, Leborgne F. Revisiting the mammographic follow-up of BI-RADS category 3 lesions. *AJR* 2002; 179:691–695
13. Vizcaíno I, Gadea L, Andreo L, et al.; Screening Program Working Group. Short-term follow-up results in 795 nonpalpable probably benign lesions detected at screening mammography. *Radiology* 2001; 219:475–483
14. Morris EA, Comstock CE, Lee CH, et al. ACR BI-RADS Magnetic Resonance Imaging. In: D'Orsi CJ, Sickles EA, Mendelson EB, et al. *ACR BI-RADS atlas, Breast Imaging Reporting and Data System*. Reston, VA: American College of Radiology, 2013
15. Chikarmane SA, Birdwell RL, Poole PS, Sippo DA, Giess CS. Characteristics, malignancy rate, and follow-up of BI-RADS category 3 lesions identified at breast mr imaging: implications for MR image interpretation and management. *Radiology* 2016; 280:707–715
16. Eby PR, DeMartini WB, Gutierrez RL, Saini MH, Peacock S, Lehman CD. Characteristics of probably benign breast MRI lesions. *AJR* 2009; 193:861–867
17. Eby PR, Demartini WB, Peacock S, Rosen EL, Lauro B, Lehman CD. Cancer yield of probably benign breast MR examinations. *J Magn Reson Imaging* 2007; 26:950–955
18. Grimm LJ, Anderson AL, Baker JA, et al. Frequency of malignancy and imaging characteristics of probably benign lesions seen at breast MRI. *AJR* 2015; 205:442–447
19. Kriege M, Brekelmans CT, Boetes C, et al.; Magnetic Resonance Imaging Screening Study Group. Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. *N Engl J Med* 2004; 351:427–437
20. Kuhl CK, Schmutzler RK, Leutner CC, et al. Breast MR imaging screening in 192 women proved or suspected to be carriers of a breast cancer susceptibility gene: preliminary results. *Radiology* 2000; 215:267–279
21. Liberman L, Morris EA, Benton CL, Abramson AF, Dershaw DD. Probably benign lesions at breast magnetic resonance imaging: preliminary experience in high-risk women. *Cancer* 2003; 98:377–388
22. Lourenco AP, Chung MT, Mainiero MB. Probably benign breast MRI lesions: frequency, lesion type, and rate of malignancy. *J Magn Reson Imaging* 2014; 39:789–794
23. Sadowski EA, Kelcz F. Frequency of malignancy in lesions classified as probably benign after dynamic contrast-enhanced breast MRI examination. *J Magn Reson Imaging* 2005; 21:556–564
24. Spick C, Szolar DH, Baltzer PA, et al. Rate of malignancy in MRI-detected probably benign (BI-RADS 3) lesions. *AJR* 2014; 202:684–689
25. Weinstein SP, Hanna LG, Gatsonis C, Schnall MD, Rosen MA, Lehman CD. Frequency of malignancy seen in probably benign lesions at contrast-enhanced breast MR imaging: findings from ACRIN 6667. *Radiology* 2010; 255:731–737
26. Strigel RM, Rollenhagen J, Burnside ES, et al. Screening breast MRI outcomes in routine clinical practice: comparison to BI-RADS benchmarks. *Acad Radiol* 2017; 24:411–417
27. Spick C, Bickel H, Polanec SH, Baltzer PA. Breast lesions classified as probably benign (BI-RADS 3) on magnetic resonance imaging: a systematic review and meta-analysis. *Eur Radiol* 2018; 28:1919–1928
28. Lee CI, Ichikawa L, Rochelle MC, et al. Breast MRI BI-RADS assessments and abnormal interpretation rates by clinical indication in US community practices. *Acad Radiol* 2014; 21:1370–1376
29. Niell BL, Gavenonis SC, Motazedi T, et al. Auditing a breast MRI practice: performance measures for screening and diagnostic breast MRI. *J Am Coll Radiol* 2014; 11:883–889
30. Sedora Román NI, Mehta TS, Sharpe RE, et al. Proposed biopsy performance benchmarks for MRI based on an audit of a large academic center. *Breast J* 2018; 24:319–324
31. Smith RA, Andrews KS, Brooks D, et al. Cancer screening in the United States, 2017: a review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin* 2017; 67:100–121
32. Panigrahi B, Harvey SC, Mullen LA, et al. Characteristics and outcomes of BI-RADS 3 lesions on breast MRI. *Clin Breast Cancer* 2019; 19:e152–e159
33. Lee AY, Ichikawa L, Lee JM, et al. Concordance of BI-RADS assessments and management recommendations for breast MRI in community practice. *AJR* 2016; 206:211–216

Study Guide

Frequency and Cancer Yield of BI-RADS Category 3 Lesions Detected at High-Risk Screening Breast MRI

Alan Mautz[1], Joseph J. Budovec[2]

[1]The Aroostook Medical Center, Presque Isle, ME.
[2]Medical College of Wisconsin, Milwaukee, WI.
*amautz@emhs.org, jbudovec@mcw.edu**

Introduction

1. What is the estimated cancer risk of BI-RADS 3 lesions at mammography? Does this estimated cancer risk translate to similar risks at breast MRI? What cancer risk or malignancy rate is currently suggested in the most recent BI-RADS atlas for breast MRI?
2. What are the clinical questions that the study attempts to answer?
3. What are the indications for breast MRI? Which of these indications is the most common indication for breast MRI?

Methods

4. What study design was used? How were imaging studies selected? Which types of imaging studies were excluded from analysis in this study and why?
5. What data were collected from the included imaging studies?
6. How were data regarding breast density for the included imaging studies collected?
7. What are the limitations of this study? Are these limitations adequately discussed?

Results

8. How often was a BI-RADS 3 assessment made? Between baseline and nonbaseline studies, when was BI-RADS 3 more often used?
9. What characteristics at breast MRI were most commonly associated with a BI-RADS 3 assessment? Of the BI-RADS 3 examinations, which breast imaging features were associated with subsequently diagnosed breast cancer?
10. What were the most common indications for the BI-RADS 3 examinations with subsequent diagnosis of malignancy?

Discussion

11. How often do you use the BI-RADS 3 assessment during breast MRI interpretation? What rate does the American College of Radiology 2013 BI-RADS atlas suggest for an established MRI screening program?
12. What caution does the study suggest regarding the use of the BI-RADS 3 assessment for nonbaseline MRI examinations?
13. The study results suggest that a 6-month follow-up window for BI-RADS 3 examinations yields the best results for detecting malignancy. How would you convey this information to clinicians in the setting of health insurance companies' possible hesitation about allowing breast MRI more often than every 12 months?

Background Reading

1. Chikarmane SA, Birdwell RL, Poole PS, Sippo DA, Giess CS. Characteristics, malignancy rate, and follow-up of BI-RADS category three lesions identified at breast MRI: implications for MR image interpretation and management. *Radiology* 2016; 280:707–715
2. Eby PR, DeMartini WB, Gutierrez RL, Saini MH, Peacock S, Lehman CD. Characteristics of probably benign breast MRI lesions. *AJR* 2009; 193:861–867
3. Monticciolo DL, Newell MS, Moy L, Niell B, Monsees B, Sickles EA. Breast cancer screening in women at higher-than-average risk: recommendations from the ACR. *J Am Coll Radiol* 2018; 15:408–414

*Please note that the authors of the Study Guide are distinct from those of the companion article.