RADIOLOGIC
CLINICS OF
NORTH AMERICA

Radiol Clin N Am 40 (2002) 409–430

Breast imaging reporting and data system (BI-RADS)

Laura Liberman, MD*, Jennifer H. Menell, MD

Department of Radiology, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA

The Breast Imaging Reporting and Data System (BI-RADS) lexicon was developed by the American College of Radiology (ACR) to standardize mammographic reporting [1–3]. The lexicon includes terms for describing breast parenchymal patterns, features of masses and calcifications, associated findings, and final assessment categories. Potential benefits of the lexicon include increased clarity in reporting, improved communication, and facilitation of research, particularly across different institutions. This article reviews the terms defined in the BI-RADS lexicon for mammography, describes strengths and limitations of the lexicon, and discusses the preliminary work relating to the development of standardized lexicons for breast sonography and breast MRI.

BI-RADS lexicon for mammography

The BI-RADS lexicon describes four classes of breast parenchymal density: class 1, almost entirely fat; class 2, scattered fibroglandular densities; class 3, heterogeneously dense; and class 4, dense (Fig. 1). A mass is defined as a space-occupying lesion seen in two different projections; if a potential mass is seen in only a single projection, it should be called a density until its three-dimensionality is confirmed. Mass margins are described as circumscribed, microlobulated (undulate in short cycles), obscured (hidden by superimposed adjacent tissue), indistinct (poor definition not caused by superimposed tissue, raising the possibility of infiltration of the lesion into adjacent tissue), and spiculated (lines radiate from the margins) (Fig. 2). Mass shape can be described as round, oval, lobular, or irregular. Architectural distortion is shape with radiating spicules but no definite mass visible (Fig. 3). Mass density can be described as high, equal, low, or fat containing.

The lexicon also defines special cases, including: intramammary lymph node (typically reniform or with radiolucent notch because of fat in the hilum, most often seen in the upper outer quadrant) (Fig. 4); solitary dilated duct (usually of minor significance unless it represents an interval change from prior mammograms); asymmetric breast tissue (judged relative to the corresponding area in the contralateral breast, usually a normal variant, but may be important when it corresponds to a palpable asymmetry); focal asymmetric density (a density that cannot be accurately described using the other shapes, could represent an island of fibroglandular tissue, but may warrant additional evaluation)

The lexicon defines specific terms to describe the shapes (morphology) of calcifications and the patterns in which they are arrayed in the breast parenchyma (distribution). Morphologic descriptors are *typically benign*, *intermediate concern*, and *higher probability of malignancy*. Typically benign calcifications include skin, vascular, coarse or popcorn-like, large rod-like, round (or punctate if smaller than 0.5 mm), lucent-centered, eggshell or rim, milk of calcium, suture, and dystrophic (Fig. 5). Intermediate concern calcifications are amorphous or indistinct; these calcifications are often round or "flake" shaped and are sufficiently small or hazy that a more specific morphologic classification cannot be determined. Calcifications with a higher probability of malignancy include pleomorphic or heterogeneous calcifications (formerly called granular) and fine linear or fine, linear, branching (casting) calcifications (Fig. 6). The distribution of

* Corresponding author.
E-mail address: libermal@mskcc.org (L. Liberman).

PII: S0033-8389(01)00017-3

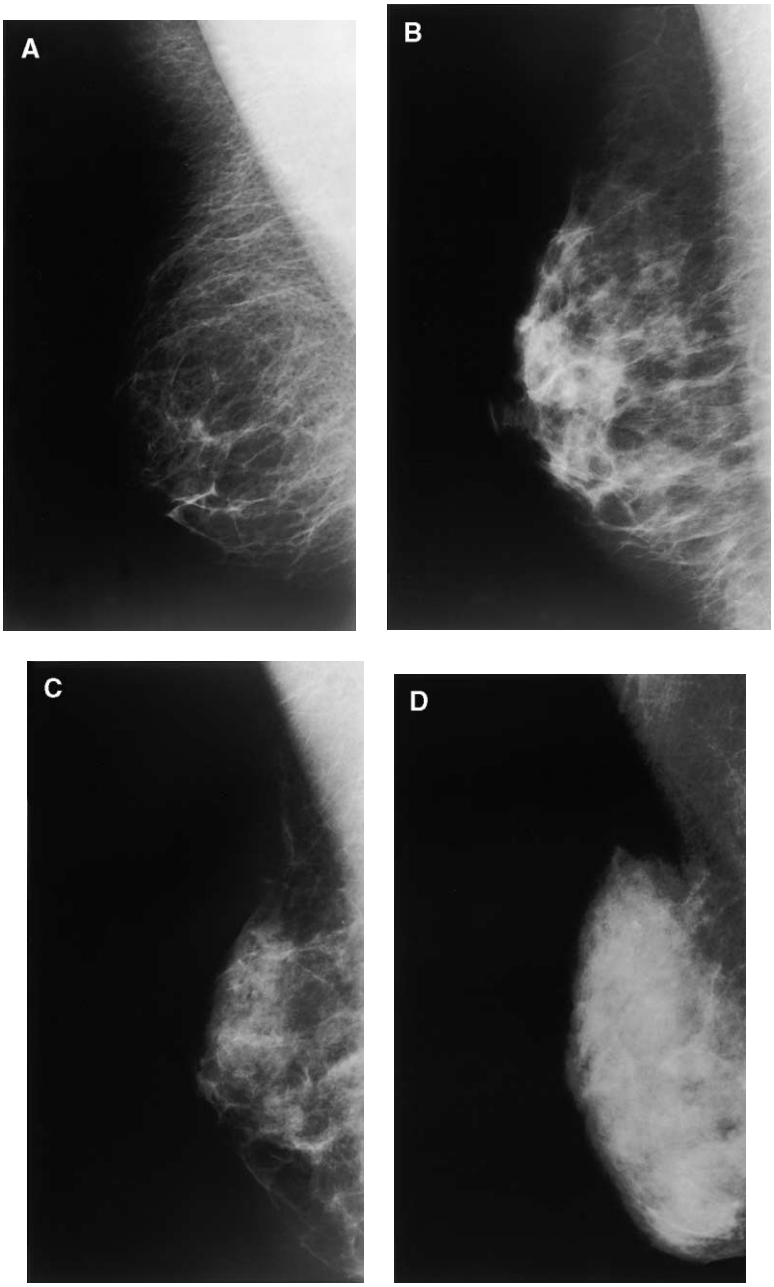

Fig. 1. Breast parenchymal density as seen on mediolateral oblique view mammograms. (A) Fatty (ACR class 1); (B) Mildly dense (ACR class 2); (C) Moderately dense (ACR class 3); (D) Dense (ACR class 4).

calcifications has been described as grouped or clustered (multiple calcifications in less than 2 mL tissue), linear, segmental (suggesting deposits in a duct), regional (large volume not necessarily conforming to a duct distribution), diffuse/scattered (random distribution), or multiple.

In addition, the lexicon defines associated findings, used with masses or calcifications or alone when no other abnormality is present, including skin or nipple retraction, skin or trabecular thickening, skin lesion, axillary adenopathy, or architectural distortion. The lexicon suggests that the location of the lesion be

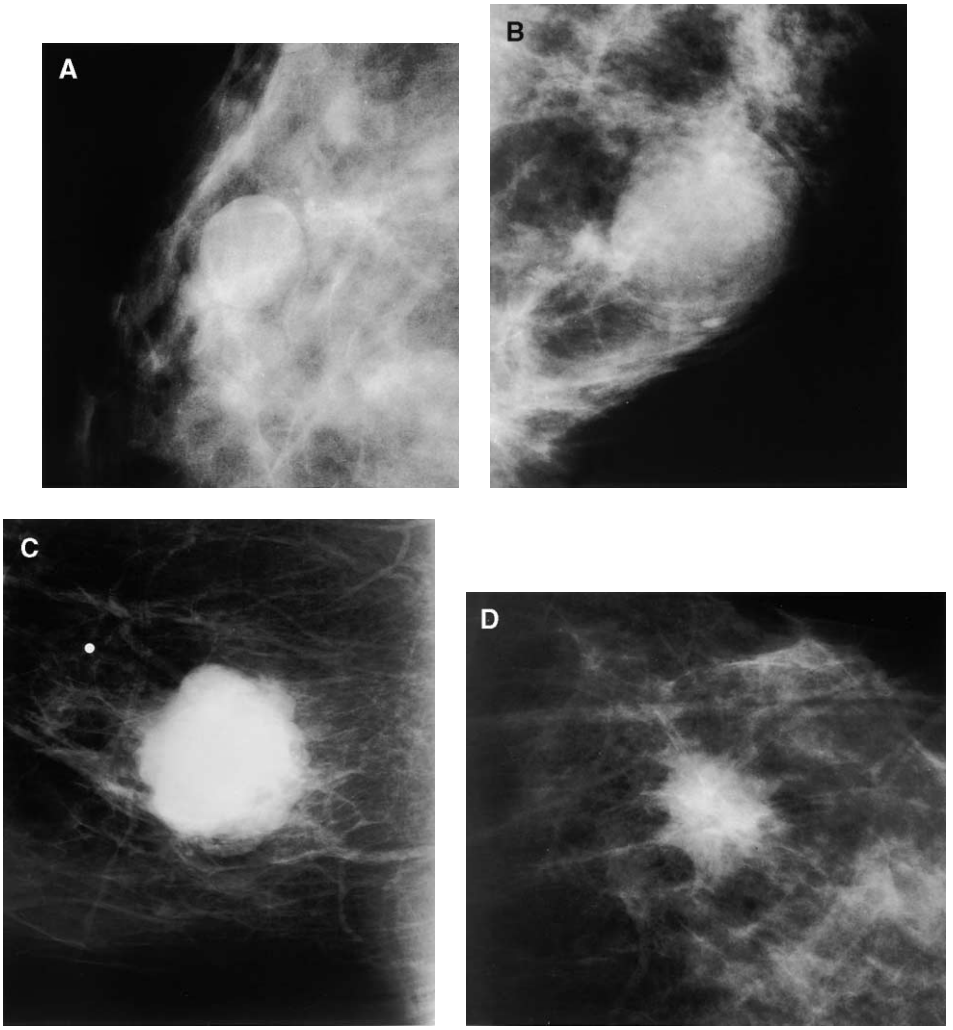

Fig. 2. Mass margin characteristics as defined by the BI-RADS lexicon. (A) Circumscribed mass, shown to be a simple cyst at sonography. (B) Partially obscured mass; sonography showed as simple cyst. (C) Microlobulated mass corresponding to palpable lump denoted by radiopaque skin marker; biopsy showed infiltrating ductal carcinoma and ductal carcinoma in situ (DCIS). (D) Spiculated mass; biopsy showed infiltrating ductal carcinoma and DCIS.

expressed by indicating the side (left, right, or both), the location (according to the face of the clock and subareolar, central, or axillary tail, if appropriate) and the depth of the lesion (anterior, middle, or posterior).

Perhaps most important, the lexicon defines assessment categories to describe the radiologist's level of suspicion regarding the mammographic finding (Table 1). As of April 1999, it has been required by law that all mammography reports in the United States contain a BI-RADS assessment category, with its description in layman's terms. Note that although there are six assessment categories, there are only four possible outcomes: additional imaging studies (category 0), routine annual mammography (category 1 or 2), 6-month follow-up (category 3), and biopsy (category 4 or 5).

Potential usefulness of the lexicon

Final assessment categories

Final assessment categories of the BI-RADS lexicon are useful predictors of malignancy. In three published series, the frequency of carcinoma was significantly higher for BI-RADS category 5 (highly

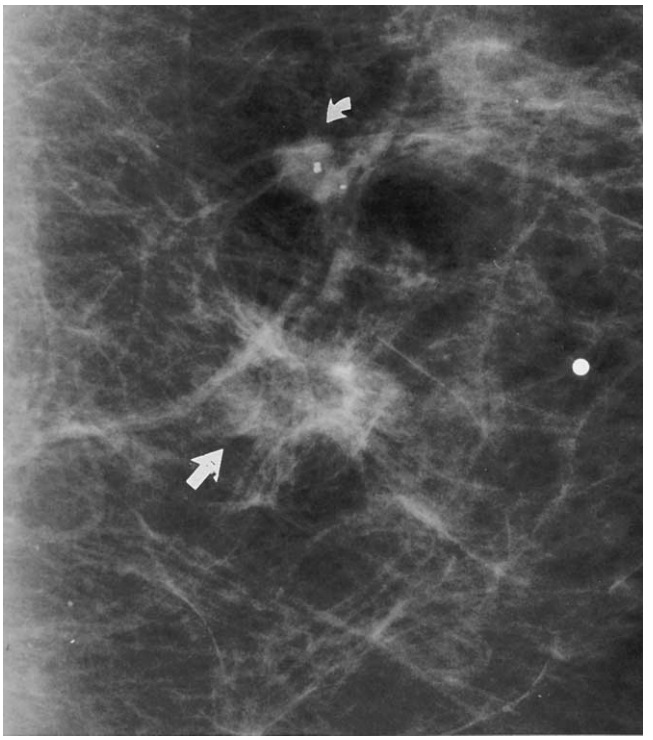

Fig. 3. Spiculated architectural distortion at mammography (*straight arrow*), corresponding to a vaguely palpable thickening denoted by radiopaque skin marker. Biopsy yielded infiltrating lobular carcinoma. There was an adjacent lobulated mass with coarse calcification (*curved arrow*), stable from prior years and consistent with a benign fibroadenoma.

suggestive of malignancy) than for category 4 (suspicious), ranging from 81% to 97% for category 5 versus 23% to 34% for BI-RADS category 4 (Table 2) [4–6]. Liberman et al [5] found a significantly higher frequency of carcinoma among category 5 than among category 4 lesions for all mammographic findings and all interpreting radiologists.

Except for some guidelines regarding calcification morphology, the lexicon does not explicitly state which mammographic features should be included in the different final assessment categories. Analysis of the descriptive terms of the lexicon, however, allows some recommendations to be made. In an analysis of 492 lesions that had needle localization and surgical biopsy, Liberman et al [5] found that the features with highest positive predictive value for masses were spiculated borders and irregular shape (Table 3). For calcifications, they were linear morphology and segmental or linear distribution (Table 4). On the basis of this finding, they recommended that these findings warrant a designation of category 5. Further study is needed to better define the mammographic patterns with the highest positive predictive value and those that have the highest likelihood of representing benign disease.

BI-RADS category 3: probably benign

A potential advantage of the lexicon is precise definition of lesions that are probably benign, allowing women with probably benign lesions the option of mammographic surveillance rather than biopsy. Few studies have addressed the frequency of a BI-RADS category 3 (probably benign) designation. Caplan et al [7] reported that 7.7% of 372,760 mammograms performed as part of the National Breast and Cervical Cancer Detection Program were classified as category 3. They found the probability of receiving a category 3 classification was higher in women who were young, symptomatic, or had abnormal findings on clinical breast examinations. They also reported that the percentage of mammograms classified as category 3 by state or tribal organization ranged from 1.4% to 14.0%, suggesting variability among radiologists in using this BI-RADS code for probably benign lesions.

Fig. 4. A benign intramammary lymph node (BI-RADS category 2). Note the notch corresponding to the fatty hilum.

Although one series in the surgical literature noted that almost half the lesions referred for biopsy were in category 3 (probably benign) [4], published studies in the radiology literature indicate that approximately 70% of lesions referred for biopsy are in category 4 and that approximately 20% are in category 5, with only a small number of category 3 lesions referred for biopsy [5,6]. Several studies published before and after introduction of the BI-RADS lexicon support the use of short-term follow-up mammography for probably benign lesions.

Sickles [8] prospectively evaluated the value of short-term follow-up mammography in 3184 patients with baseline mammographic lesions classified as probably benign in a study published before the BI-RADS lexicon. Lesions were only classified as probably benign after careful evaluation, including magnification images. All probably benign lesions were evaluated with a short-term follow-up mammography protocol that included imaging the ipsilateral breast 6 months after the initial mammogram, and then both breasts 12, 24, and 36 months after the initial mammogram, to document stability.

Of the 3184 probably benign lesions included in the study, cancer was subsequently discovered in 17 (0.5%) [8]. Fifteen of the 17 cancers were diagnosed by means of interval change at follow-up mammography before they were palpable; all 17 were stage 0 or stage I at the time of diagnosis (one positive axillary lymph node was present in two patients; one had a circumscribed solid nodule and one had an asymmetric area of fibroglandular tissue). Cancer was discovered in 1 of 1234 (0.1%) clusters of round or punctate calcifications, 12 of 589 (2%) solitary solid circumscribed masses, 2 of 448 (0.4%) focal asymmetric densities, 1 of 522 (0.2%) scattered or randomly clustered calcifications, and 1 of 253 (0.4%) multiple solid circumscribed nodules.

Sickles [9] has also addressed the question of whether patient age or lesion size should prompt immediate biopsy of nonpalpable, circumscribed, solid nodules. Of 1403 cases included in this study, cancer was found in 19 (1.4%). Only small differences in the frequency of cancer were found for various patient age and lesion size subgroups. Even in the group of women aged 50 and older, the frequency of cancer was 1 of 560 (1.7%). These data suggest that lesion size and patient age should not deter from recommending short-interval follow-up mammography for nonpalpable circumscribed solid masses.

A second large-scale prospective study evaluating the use of short-term follow-up for probably benign lesions was published before the BI-RADS lexicon

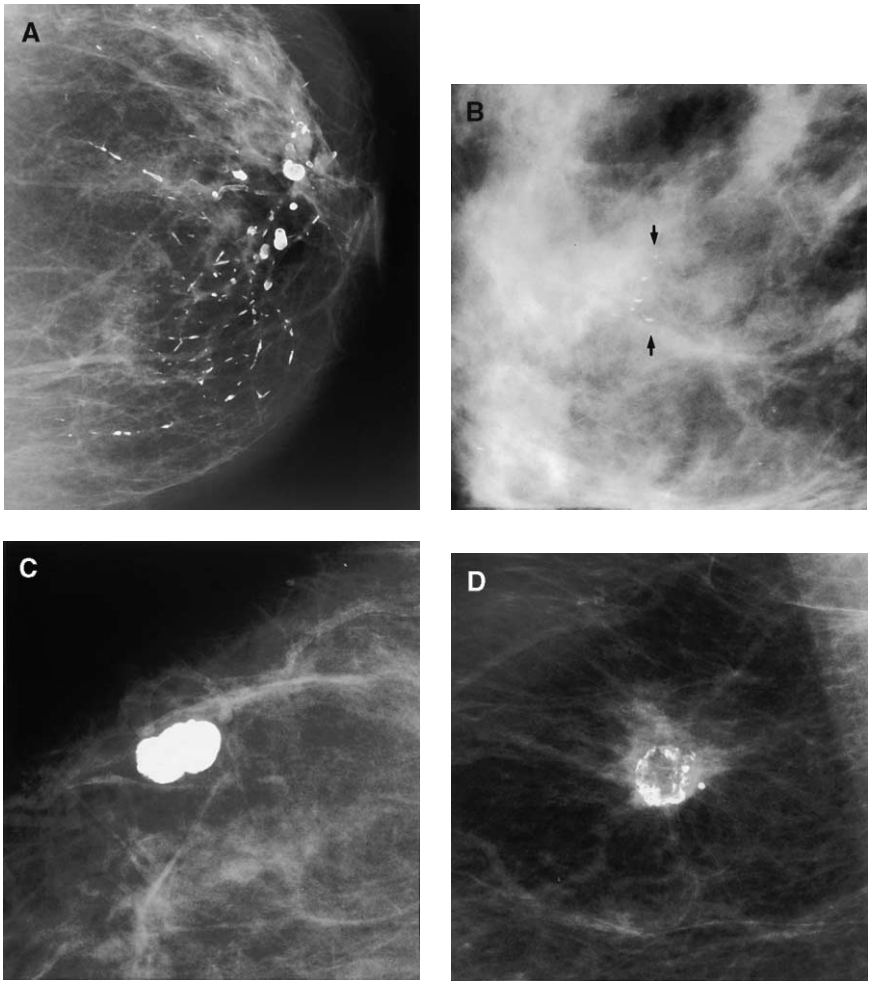

Fig. 5. Typically benign calcifications. (A) Variety of benign calcifications: peripherally calcified oil cysts of fat necrosis, large rod-like calcifications of secretory disease, and vascular calcifications. (B) Milk of calcium. Note the layering or "teacup" appearance of this 90° lateral magnification view (*arrows*). (C) Popcorn calcification typical of fibroadenoma. (D) Eggshell calcifications associated with architectural distortion in area of postoperative fat necrosis.

by Varas et al [10]. Probably benign lesions in this study included single or multiple circumscribed masses, multiple rounded, clustered, or scattered calcifications within less than one quadrant of the breast, and abnormal parenchymal opacities (areas of localized dense tissue, without definable margins or architectural distortion, identified on two views). Carcinoma was found in 9 of 535 (1.7%) probably benign lesions, including 4 of 289 (1.4%) solitary circumscribed masses, 4 of 104 (3.8%) lesions evident as microcalcifications, and 1 of 54 (1.9%) abnormal parenchymal opacities. Of the nine carcinomas identified, two were ductal carcinoma in situ (DCIS) and seven were invasive carcinomas (including one DCIS with microinvasion); two had positive axillary nodes. These data also support the use of short-term follow-up as an alternative to biopsy for probably benign (BI-RADS category 3) lesions.

If short-term follow-up is selected, interval progression (increase in size of a mass or increase in number of calcifications) at follow-up should prompt a biopsy. In Sickles' [8] study, carcinoma was identified in 15 of 131 (11%) biopsies performed for mammographic progression; in the study of Varas et al [10], 9 of 16 (56%) lesions that demonstrated mammographic progression were found to represent carcinoma. In both studies, no carcinomas were identified in probably benign lesions that remained

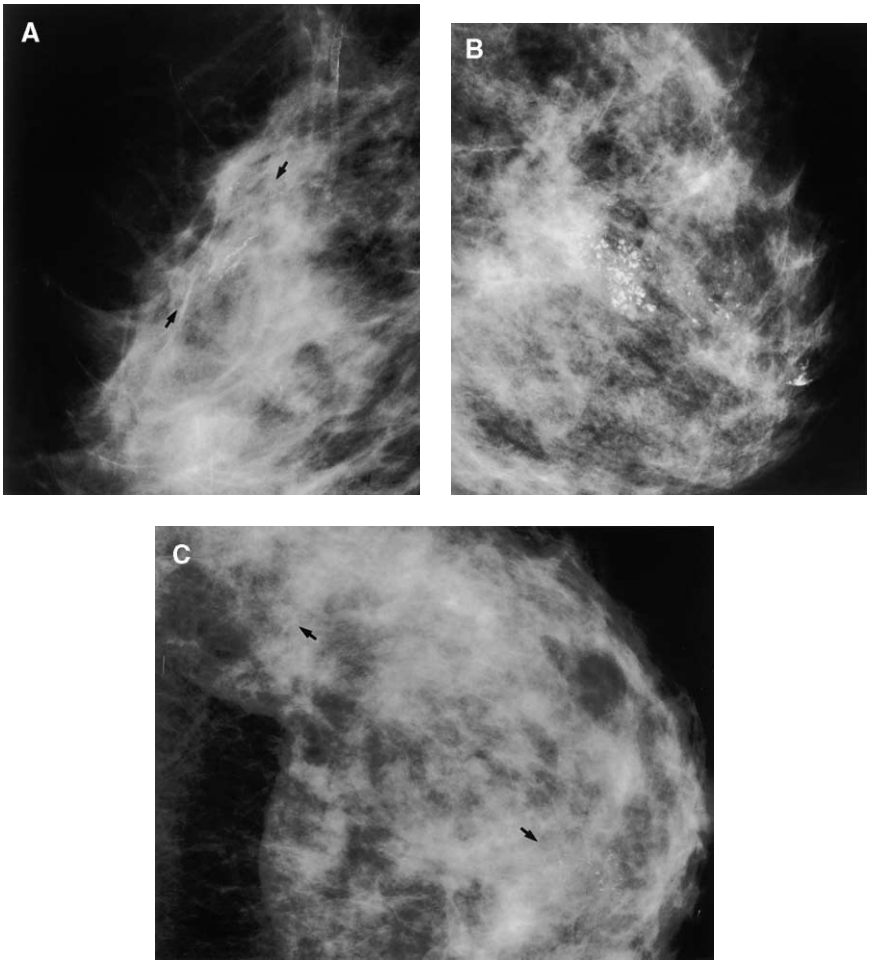

Fig. 6. Calcifications with higher probability of malignancy. (A) Calcifications with linear morphology and linear distribution (*arrows*). Biopsy yielded ductal carcinoma in situ (DCIS) with calcification. (B) Pleomorphic calcifications in segmental distribution. Biopsy yielded infiltrating ductal carcinoma and with calcifications present in DCIS. (C) Two clusters of pleomorphic calcifications (*arrows*). Both yielded DCIS with calcifications at biopsy, and the patient was treated with mastectomy.

stable on follow-up mammography. Careful attention to the follow-up protocol should allow us to detect carcinoma at an early stage while minimizing the number of benign biopsies.

In an update of data from the University of California at San Francisco, Sickles [11] noted that the frequency of cancer among probably benign lesions was 0.7% (33 of 4533), with the likelihood of malignancy 23 of 1692 (1.4%) for solid circumscribed masses, 3 of 502 (0.6%) for focal asymmetric densities, 5 of 1338 (0.4%) for localized microcalcifications, 1 of 329 (0.3%) for multiple circumscribed masses, 1 of 619 (0.2%) for generalized microcalcifications, and 0 (0%) for other miscellaneous findings. With further update to 7484 probably benign lesions, Sickles [11] reported carcinoma in 36 (0.5%). Of these 36 cancers found at periodic mammographic surveillance, 6 (16.7%) were identified at the 6-month follow-up mammogram, 2 (5.6%) by palpation between 6 months and 1 year, 17 (47.2%) at the 1-year mammogram, 2 (5.6%) by palpation between year 1 and year 2, 7 (19.4%) at the 2-year follow-up mammogram, and 2 (5.6%) at the 3-year follow-up mammogram. Thirty-five (97%) of these 36 cancers were smaller than 2 cm at diagnosis, and 34 (94.4%) were node-negative at the time of diagnosis; two each had one positive node; none had distant metastases.

The potential benefits of short-term follow-up mammography for probably benign lesions were recently restated by Sickles [11]. He noted that 95%

Table 1
Assessment categories of the BI-RADS lexicon

Stage	Result
0	Assessment incomplete. Need additional imaging evaluation.
1	Negative. Routine mammogram in 1 year recommended.
2	Benign finding. Routine mammogram in 1 year recommended.
3	Probably benign finding. Short-interval follow-up suggested.
4	Suspicious. Biopsy should be considered.
5	Highly suggestive of malignancy. Appropriate action should be taken.

Data from American College of Radiology. Breast Imaging Reporting and Data System (BI-RADS). Reston, VA: College of Radiology; 1995; with permission.

of patients complied with at least half of the recommended examinations in the follow-up protocol, and 50% completed the entire protocol. He also noted that only approximately 2% of women chose biopsy rather than follow-up. Compared to percutaneous core biopsy, follow-up lowers the cost by a factor of 8, with savings of $1040 per probably benign lesion; it is also associated with lower patient stress. Although existing data support that probably benign lesions can be identified and safely managed with short-term follow-up mammography, the management of BI-RADS category 3 lesions continues to be debated [12].

Breast parenchymal density

Literature before the BI-RADS lexicon defined different breast density parenchymal patterns and evaluated the frequency of carcinoma among women with different breast densities [13–15]. Analysis of the impact of breast density on breast cancer incidence are complicated by the inverse relationship between age and breast parenchymal density and by the lower sensitivity of mammography in women with dense breasts. The BI-RADS lexicon potentially allows standardization of reporting of breast parenchymal density, facilitating further research in this area.

Dense breast tissue interferes with interpretation of mammograms. Mandelson et al [16] evaluated breast density as a predictor of mammographic detection. Mammographic sensitivity was 80% among women with predominantly fatty breasts (ACR class 1) but 30% in women with extremely dense breasts (ACR class 4). The odds ratio for interval cancer among women with extremely dense breasts was 6.14 (95% confidence interval [CI], 1.95–19.4), compared with women with extremely fatty breasts, after adjustment for age at index mammogram, menopausal status, use of hormone replacement therapy, and body mass index. When only those interval cancer cases confirmed by retrospective review of index mammograms were considered, the odds ratio rose to 9.47 (95% CI, 2.78–32.3).

Although it remains controversial, it has been suggested that mammographic density may be an independent risk factor for development of breast cancer. Satija et al [17] reviewed results of 82,391 screening mammograms among 36,495 women aged 40 to 80 with no history of breast cancer. They found that ACR class 1 and 2 breasts, at age 40, were associated with a relative risk of 0.39 with respect to the general population at the same age, whereas at age 80 the relative risk was 0.61. The relative risk for ACR class 3 was 0.72 at age 40

Table 2
Final assessment categories: number of lesions referred for biopsy and positive predictive value

	BI-RADS category		
Investigator	3	4	5
No. lesions referred for biopsy			
Liberman [5]	8/492 (2)	355/492 (72)	129/492 (26)
Orel [6]	141/1312 (11)	936/1312 (71)	170/1312 (13)
Lacquement [4]	322/688 (47)	234/688 (34)	106/688 (15)
PPV			
Liberman [5]	0/8 (0)	120/355 (34)	105/129 (81)
Orel [6]	3/141 (2)	279/936 (30)	165/170 (97)
Lacquement [4]	9/322 (3)	54/234 (23)	97/106 (92)

Numbers in parentheses are percentages.
PPV = positive predictive value, which is equal to the number of cancers divided by total number of lesions that underwent biopsy in that category.

Table 3
Frequency of carcinoma versus combinations of features: mass shape and margins

	Mass shape					
Mass margins	Irregular[a]	Round	Lobulated	Oval	Distortion	Total
Spiculated[b]	45/54 (83)	6/6 (100)	—	1/1 (100)	4/8 (50)	56/69 (81)
Indistinct	20/35 (57)	5/14 (36)	3/9 (33)	1/8 (13)	—	29/66 (44)
Obscured	—	2/3 (67)	1/3 (33)	0/3 (0)	—	3/9 (33)
Microlobulated	—	0/2 (0)	1/2 (50)	0/2 (0)	—	1/6 (17)
Circumscribed	1/1 (100)	0/6 (0)	1/4 (25)	0/11 (0)	—	2/22 (9)
Total	66/90 (73)	13/31 (42)	6/18 (33)	2/25 (8)	4/8 (50)	91/172 (53)

Data refer to lesions that were subject to surgical biopsy. Numbers in parentheses are percentages. Dash (—) indicates there were no lesions with the specified combination of features.
Reprinted from Liberman L, Abramson AF, Squires FB, Glassman J, Morris EA, Dershaw DP. The Breast Imaging Reporting and Data System: positive predictive value of mammographic features and final assessment categories. AJR Am J Roentgenol 1998;171:35–40; with permission.

[a] Frequency of carcinoma was significantly higher for spiculated margins than for all other margin characteristics (56/69 = 81% versus 35/103 = 34%, $P < 0.001$, relative risk 2.4 [95% confidence intervals 1.8–3.2]).

[b] Frequency of carcinoma was significantly higher for irregular shape than for all other shapes (66/90 = 73% versus 25/82 = 30%, $P < 0.001$, relative risk 2.4 [95% confidence intervals 1.7–3.4]).

and 1.13 at age 80. ACR class 4 was divided into two groups with respect to risk, with the relative risk for the densest pattern as high as 2.49 times the risk of the patterns in the general population. Additional study is necessary to further evaluate the impact of breast density on mammographic interpretation and breast cancer incidence and to assess the use of computer-aided diagnostic techniques in quantifying parenchymal density and its associated risk.

Computer-aided diagnosis

It has been suggested that computer-aided diagnostic techniques may assist in mammographic interpretation, for lesion detection and for classification. In particular, some investigators have proposed the use of an artificial neural network (ANN), a form of artificial intelligence that can be trained to "learn" essential information from a data set, may improve the positive predictive value (PPV) of

Table 4
Frequency of carcinoma versus combination of features: calcification distribution and morphology

	Calcification morphology					
Calcification distribution	Linear[a]	Pleomorphic	Amorphous	Punctate	Coarse	Total
Segmental[b]	10/10 (100)	7/12 (58)	0/1 (0)	—	—	17/23 (74)
Linear[b]	6/8 (75)	7/9 (78)	—	0/2 (0)	—	13/19 (68)
Multiple	1/1 (100)	4/6 (67)	0/2 (0)	—	—	5/9 (56)
Regional	0/1 (0)	4/9 (44)	2/3 (67)	—	—	6/13 (46)
Clustered	9/12 (75)	76/204[c] (37)	7/29 (24)	1/9 (11)	0/1 (0)	93/255 (36)
Diffuse	—	0/1 (0)	—	—	—	0/1 (0)
Total	26/32 (81)	98/241 (41)	9/35 (26)	1/11 (9)	0/1 (0)	134/320 (42)

Data refer to lesions that were subject to surgical biopsy. Numbers in parentheses are percentages. Dash (—) indicates there were no lesions with the specified combination of features.
Reprinted from Liberman L, Abramson AF, Squires FB, Glassman J, Morris EA, Dershaw DD. The Breast Imaging Reporting and Data System: positive predictive value of mammographic features and final assessment categories. AJR Am J Roentgenol 1998;171:35–40; with permission.

[a] Frequency of carcinoma was significantly higher for linear morphology than for all other morphologies (26/32 = 81% versus 108/288 = 38%, $P < 0.001$, relative risk 2.2 [95% confidence intervals 1.8–2.8]).

[b] Frequency of carcinoma was significantly higher for segmental or linear distribution than for all other distributions (30/42 = 71% versus 104/278 = 37%, $P < 0.001$, relative risk 1.9 [95% confidence intervals 1.5–2.4]).

[c] Of 320 calcification lesions that underwent surgical biopsy in this study, 204 (64%) were described as clusters of pleomorphic calcifications.

biopsy recommendations. Previous work in this area was limited by lack of standardization of terminology, diminishing the potential applicability of a common artificial neural network to multiple institutions. By providing standardized terminology, the BI-RADS lexicon may facilitate progress in computer-aided diagnostic techniques.

Baker et al [18] constructed an artificial neural network based on the BI-RADS lexicon. Eighteen inputs to the network included 10 BI-RADS lesion descriptors and eight input values from the patient's medical history. The network was trained and tested on 206 cases, of which 73 were malignant. They found that at a specified output threshold, the ANN would have improved the PPV of biopsy from 35% to 61%, with a relative sensitivity of 100%. At a fixed sensitivity of 95%, the specificity of the ANN (62%) was significantly higher than that of the radiologists (30%) ($P < 0.01$). These data suggest that the BI-RADS lexicon provides a standardized language between mammographers and an ANN that can improve the PPV of breast biopsy.

In a subsequent study, Baker et al [19] studied the performance and interobserver and intraobserver variability of an artificial neural network for predicting breast biopsy outcome. Five radiologists used the BI-RADS terminology to describe 60 mammographically detected lesions, including 23 cancers. Interobserver and intraobserver variability were evaluated with the κ statistic. They found that the ANN maintained 100% sensitivity while improving the PPV of biopsy from 38% (23 of 60) to between 58% (23 of 40) and 66% (23 of 35; $P < 0.001$). Interobserver variability for interpretation of the lesions was significantly reduced by the ANN ($P < 0.001$); there was no statistically significant effect on nearly perfect intraobserver reproducibility. The authors concluded that use of an ANN with radiologists' descriptions of abnormal findings might improve the interpretation of mammographic abnormalities.

Limitations of the lexicon

Interobserver and intraobserver variability

The issue of variability in mammographic interpretation has been a subject of intense scrutiny. Elmore et al [20] published a study in which 10 radiologists reviewed 150 mammograms, including 27 in women with breast cancer. Immediate work-up was recommended for 74% to 96% of women with cancer and 11% to 65% of women without cancer. Beam et al [21] reported results of 108 radiologists who reviewed screening mammograms from 79 women, 45 of whom had cancer. Screening sensitivities ranged from 47% to 100%, and specificity ranged from 36% to 99%. The wide variation noted in these studies may be multifactorial, likely reflecting differences in detection, intervention threshold, and inclusion of subtle cases [22,23]. Reduction of interobserver and intraobserver variability is a potential benefit of the BI-RADS lexicon.

Observer variability in the use of the BI-RADS lexicon was first evaluated by Baker et al [24]. In that study, 60 mammograms were evaluated independently by five radiologists; one radiologist read each case twice. Readers were asked to select a single term from the BI-RADS lexicon for a variety of lesion descriptors. Interobserver and intraobserver variability was assessed by means of the κ statistic, with $\kappa \leq 0.2$ indicating slight agreement; $\kappa = 0.21–0.4$, fair agreement; $\kappa = 0.41–0.6$, moderate agreement; $\kappa = 0.61–0.8$, substantial agreement; and $\kappa = 0.81–1.0$, almost perfect agreement. Baker et al [24] noted substantial agreement between readers for choosing terms to describe masses and calcifications and similar intraobserver agreement (Table 5). Considerable interobserver and intraobserver variabilities were noted for associated findings and special cases. Use of terms to describe calcifications did not always conform to BI-RADS–defined levels of suspicion.

Variability in mammographic interpretation has also been assessed by Kerlikowske et al in a study of 2616 mammograms, including 267 (10.2%) with cancer, with agreement assessed using the κ statistic (Table 5). They found moderate agreement between the two radiologist readers in reporting the presence of a finding when cancer was present ($\kappa = 0.54$) and substantial agreement when cancer was not present ($\kappa = 0.62$). Agreement was moderate in assigning one of the five assessment categories but was significantly lower when cancer was present relative to when cancer was not present ($\kappa = 0.46$ vs 0.56; $P = 0.02$). Agreement for reporting the presence of a finding and mammographic assessment was 2-fold more likely for examinations with less dense breasts. Intraobserver agreement in final assessment (86%, $\kappa = 0.73$) was higher than interobserver agreement (78%, $\kappa = 0.58$).

Berg et al [26] analyzed interobserver and intraobserver variability in use of BI-RADS terminology. Five experienced mammographers used the lexicon to describe and assess 103 screening mammograms, of which 30 (29%) showed cancer, and a subset of 86 diagnostic mammograms, including 23 (27%) that showed cancer. A subset of 13 mammograms was reviewed by each radiologist 2 months later. Agreement, as measured by the κ statistic, showed a wide

Table 5
Inter- and intraobserver variability in use of the BI-RADS lexicon

	Investigator					
	Baker [24]		Kerlikowske [25]		Berg [26]	
Feature	Inter	Intra	Inter	Intra	Inter	Intra
Calcifications						
Distribution	0.77	0.80	0.46	—	0.47	—
Number	0.77	0.84	—	—	—	—
Description	0.50	0.57	0.33	—	0.36	—
Masses						
Margin	0.63	0.66	0.58	—	0.40	—
Shape	0.65	0.72	0.40	—	0.28	—
Density	0.62	0.63	0.23	—	0.40	—
Other findings						
Associated	0.32	−0.02	—	—	—	—
Special cases	0.16	0.38	—	—	0.38–1.0	—
Location of primary finding	—	—	0.69	—	—	—
Finding/no finding	—	—	0.66	0.79	—	—
Primary finding	—	—	0.56	0.71	0.75	—
Breast density	—	—	0.59	0.72	0.43	—
Assessment category	—	—	0.58	0.73	0.37	0.6 (0.35–1.0)
Recommendation	—	—	0.59	0.59	—	—

Data reflect the κ statistic, with <0.2 indicating slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.0, almost perfect agreement.

range (Table 5). Lesion management was highly variable: when assessments were grouped as to whether the lesion needed immediate evaluation (BI-RADS 0, 4, or 5) versus follow-up (BI-RADS 1, 2, and 3), five observers agreed on management for only 47 (55%) of 86 lesions. Intraobserver agreement on management was seen in 47 (85%) of 55 interpretations. The authors noted that in spite of the variability, the performance of the radiologists was outstanding, with recommendations for additional evaluation or biopsy in 90% to 97% of cancers on screening and 91% to 96% on diagnostic evaluation.

The impact of training in BI-RADS on reader agreement in feature analysis was evaluated by Berg et al [27]. They developed a test set of mammograms with 54 proven lesions (28 masses and 26 calcification lesions), of which 19 (35.2%) were malignant. Twenty-seven physicians reviewed the mammograms before and after a 1-day training session in BI-RADS. Readers were asked to describe mass borders, calcification morphology, and calcification distribution, and agreement with expert consensus was assessed using the κ statistic. For mass borders, mean κ was 0.42 before training and 0.47 afterward; for microcalcification morphology, mean κ was 0.40 before training and 0.46 afterward; for microcalcification distribution, mean κ was 0.32 before training and 0.42 afterward. They concluded that after 1-day training in BI-RADS, agreement with expert consensus improved, but only moderate agreement on feature analysis was achieved.

These studies indicate that even in the presence of a standardized lexicon, variability in mammographic reporting persists. Although variability is inherent in the practice of medicine (as in all endeavors in life), some of the observed variability may reflect weakness in the lexicon itself, deficiencies in radiologist training, and differences in performance level among the different physicians. The studies identified some specific areas that may need clarification, such as "punctate" calcifications, associated findings, and special cases. A larger illustrated lexicon, currently under development, may be useful. The BI-RADS lexicon remains a work in progress and may be modified on the basis of user input and continued research.

Communication with referring clinicians

The level of understanding of BI-RADS final assessment categories by referring clinicians was recently evaluated by Vitiello et al [28]. Of 86

clinicians who responded to a survey, 46% were not aware that radiologists were required to report mammograms using BI-RADS terminology, 64% had no information or education regarding BI-RADS, and only 35% were comfortable with reports using BI-RADS. For patients with a BI-RADS category 3 (probably benign) reading, 93% of clinicians followed the radiologist's recommendation for short-term radiologic follow-up; in addition, 62% of clinicians sent BI-RADS 3 patients for further work-up, including physical examination in their offices, surgical consultation, or both. These results indicate that many referring clinicians have little knowledge of BI-RADS and are not comfortable with it. If the goal of improving communication is to be achieved, further education is needed.

Toward a lexicon for breast sonography

Lesion characterization

The classic teaching has been that breast sonography can provide excellent differentiation of cystic (Fig. 7) from solid (Fig. 8) masses but that it is of limited usefulness in distinguishing benign from malignant solid masses in the breast. Data from Stavros et al [29] challenge this paradigm.

Stavros et al [29] published results of 750 sonographically solid breast nodules that were prospectively classified as benign, indeterminate, or malignant (Fig. 9). They defined specific features they considered malignant (Table 6) and other specific features they considered benign (Table 7). If a single malignant feature was present, the nodule was excluded from the benign classification. If one of the three combinations of benign characteristics was found (Table 7), the lesion was classified as benign. If no malignant features were found and none of the combinations of benign characteristics was present, the lesion was classified as indeterminate. In 1 of 5 groups, mammograms were also classified as negative, probably benign, indeterminate, probably malignant, and malignant, a classification that preceded the BI-RADS lexicon. All lesions underwent biopsy.

Of the 750 nodules, 625 (83%) were benign and 125 (17%) were malignant. The sonographic classification had a sensitivity of 98.4% (123 of 125), specificity of 67.8% (424 of 625), positive predictive value of 38.0% (123 of 324), negative predictive value of 99.5% (424 of 426), and accuracy of

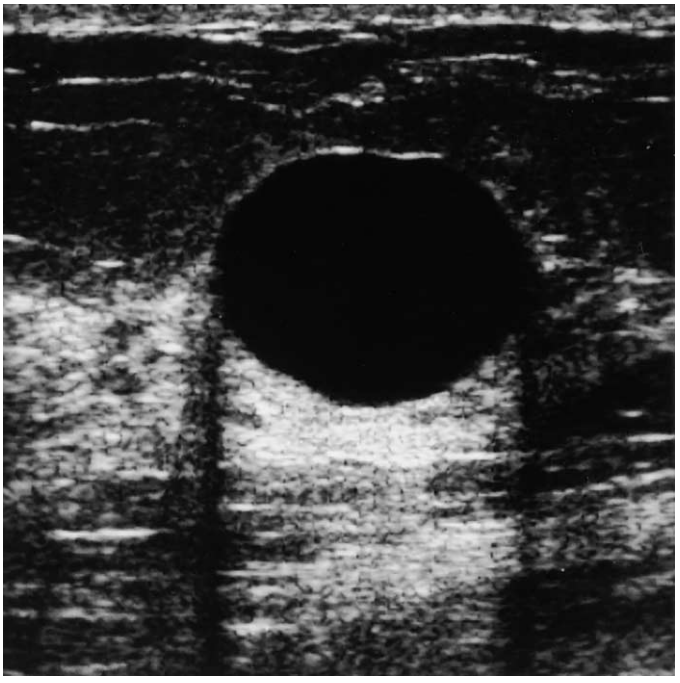

Fig. 7. Sonography of a simple cyst. Characteristics of a simply cyst include a thin wall, no internal echoes, round/oval shape, and posterior acoustic enhancement.

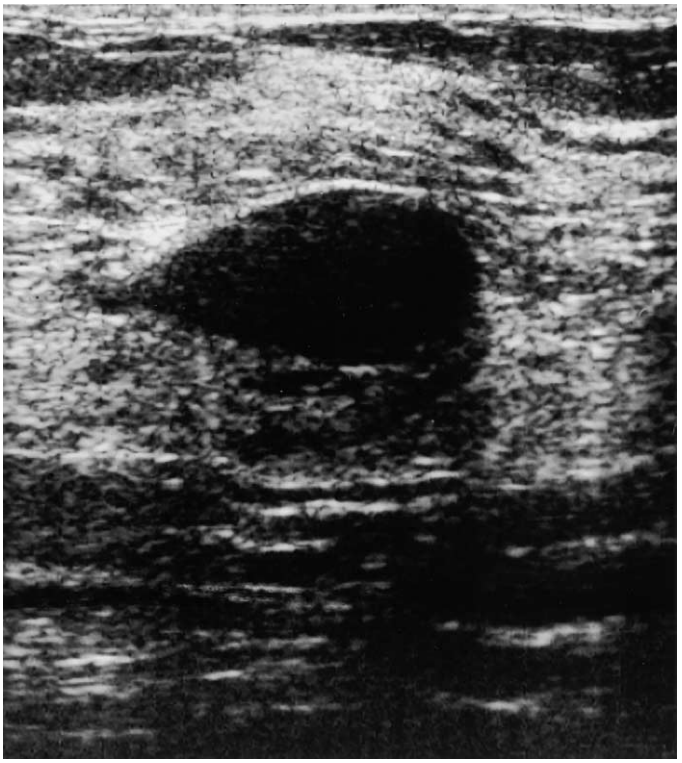

Fig. 8. Biopsy-proven fibroadenoma at sonography. Note the circumscribed borders, oval shape that is wider than it is tall, and echogenic capsule. Minimal posterior acoustic shadowing is present.

72.9% (547 of 750). Of particular interest is the negative predictive value of 99.5%. This indicates that of lesions classified as benign by sonographic criteria, only 0.5% were cancer; note that this is identical to the frequency of cancer among probably benign (BI-RADS category 3) lesions in the study by Sickles [8]. These data suggest that ultrasound may help identify lesions that have an overwhelmingly high likelihood of benignity and can be safely evaluated with short-term follow-up imaging.

The study of Stavros et al [29] also indicates that sonography can increase the radiologist's level of suspicion for lesions that prove to be cancer. Among 125 cancers, 64 (51.2%) were classified as benign (n = 20) or indeterminate (n = 44) by mammography but malignant by sonography. Among 44 palpable cancers, 32 (72.3%) were classified as benign (n = 16) or indeterminate (n = 16) by mammography but malignant by sonography.

Berg et al [30] correlated sonographic features with risk of malignancy in 588 lesions that underwent biopsy in the Radiologic Diagnostic Oncology Group V study, of which 116 (20%) were malignant. The shape feature most predictive of malignancy was irregular, with PPV of 65% for irregular, 13% for lobular, 12% for round, and 8% for oval masses. The posterior attenuation feature most predictive of malignancy was shadowing, present in half the malignant lesions; PPV was 32% for shadowing, 15% for no posterior characteristics, and 8% for posterior acoustic enhancement. Malignancy was present in 34% of lesions that had heterogeneous echotexture without cysts, 14% of homogeneous lesions, and 13% of heterogeneous lesions with cysts. Echogenicity did not discriminate between benign and malignant lesions, with PPV of 21%, 18%, and 9% for hypoechoic, hyperechoic, and isoechoic lesions, respectively. These data lend further support to the role of sonography in lesion characterization and help provide a scientific basis for the development of a BI-RADS lexicon for ultrasound.

Lexicon development

The ACR has developed an initial draft of a breast ultrasound lexicon [31], supported by the Office on Women's Health, Department of Health

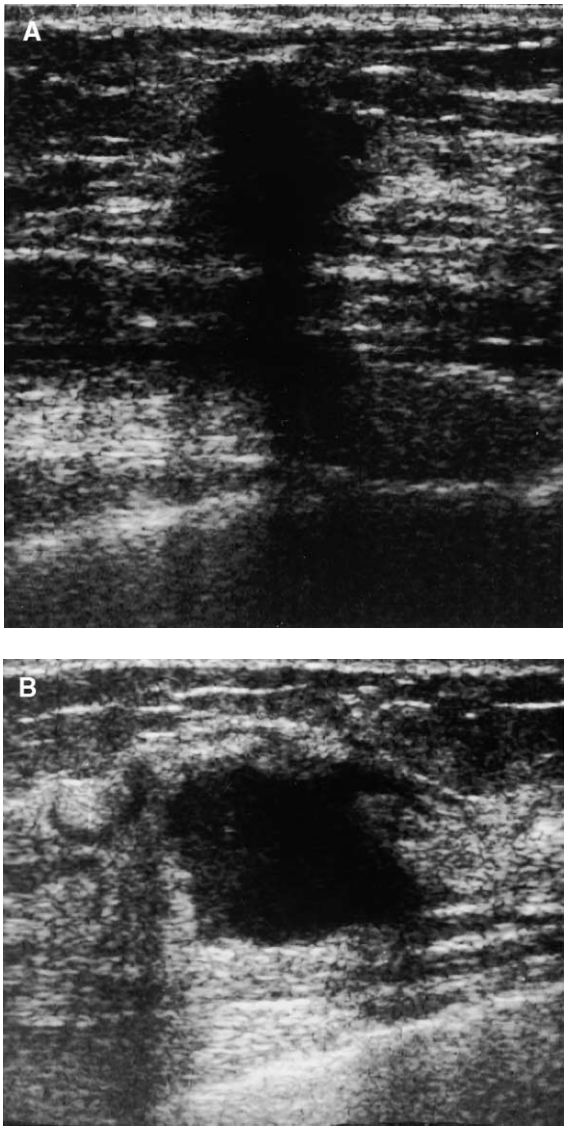

Fig. 9. Sonographic findings in breast cancers. (A) Sonography shows a spiculated, irregular, hypoechoic mass that is taller than wide and has posterior acoustic shadowing. Biopsy showed infiltrating ductal carcinoma and ductal carcinoma in situ (DCIS). (B) Sonography shows lobulated, hypoechoic solid mass with ductal extension. Posterior acoustic enhancement (a feature more common in benign lesions) is observed. Biopsy showed infiltrating ductal carcinoma and DCIS.

and Human Services. The initial draft includes descriptors for mass shape (oval, round, or irregular), echopattern (anechoic, hyperechoic, complex, or hypoechoic), and posterior acoustic features (none, enhancement, shadowing, or combined). Mass orientation is described as parallel (oriented along skin line, ''wider than tall'') or not parallel (axis not oriented along skin line, or ''taller than wide''). Mass margins are circumscribed (with no rim, thin rim, or thick rim) or irregular (indistinct, angular, microlobulated, or spiculated).

Effect on surrounding tissue is also noted, including effect on ducts or Cooper ligaments, edema, architectural distortion, skin thickening or retraction, and unclear plane with pectoral muscle. Also included are descriptors for associated calcifications (none,

Table 6
Malignant sonographic characteristics versus malignant histologic findings

Characteristics	Sensitivity	Specificity	PPV	NPV	Accuracy	OR
Spiculation	36.0	99.4	91.8	88.6	88.8	5.5
Taller than wide	41.6	98.1	81.2	89.4	88.7	4.9
Angular margins	83.2	92.0	67.5	96.5	90.5	4.0
Shadowing	48.8	94.7	64.9	90.2	87.1	3.9
Branch pattern	29.6	96.6	64.0	87.3	85.5	3.8
Hypoechogenicity	68.8	90.1	60.1	93.6	87.2	3.6
Calcifications	27.2	96.3	59.6	86.9	84.8	3.6
Duct extension	24.8	95.2	50.8	86.4	79.3	3.0
Microlobulation	75.2	83.8	48.2	94.4	82.4	2.9

Numbers reflect percentages.
PPV = positive predictive value; NPV = negative predictive value; OR = odds ratio.
Adapted from Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 1995;196:123–34; with permission.

macrocalcifications, microcalcifications in mass, microcalcifications outside of mass), special cases (mass in or on skin, foreign body, intramammary lymph nodes, or axillary lymph nodes), vascularity (cannot assess, none, same as normal tissue, decreased, or increased), and final assessment categories.

Mendelson et al [31] suggest that descriptors should be based on multiple views of masses obtained in orthogonal imaging planes and that the location of the abnormality be described using a quadrant, clock-face location, or labeled diagram of the breast, ideally including distance from the nipple. Development of a sonographic lexicon is made more complex by additional variables in sonography, including the high level of operator dependence, technical differences dependent on equipment, and availability of real-time assessment. Further work is needed to validate the lexicon terminology and to assess the positive and negative predictive values of the different descriptors.

Breast sonography: observer variability in lesion description and assessment

Baker et al [32] evaluated 60 consecutive sonographic studies of solid breast lesions. Static sonographic images of each solid breast lesion were acquired and reviewed by five radiologists experienced in breast imaging, and radiologists described mass shape, margin, echogenicity, presence of a pseudocapsule, acoustic transmission, and echotexture according to terms defined by Stavros et al [29]. Interobserver and intraobserver variability were assessed using the κ statistic (Table 8).

In that study, Baker et al [32] reported moderate interobserver agreement and substantial intraobserver agreement for most categories (Table 8). Interobserver agreement ranged from lowest for determining the presence of an echogenic capsule to highest for mass shape; intraobserver agreement was lowest for mass echotexture and highest for

Table 7
Benign sonographic characteristics versus benign histologic findings

Characteristic	Sensitivity	Specificity	PPV	NPV	Accuracy	OR
Hyperechogenicity	100.0	7.4	17.8	100.0	22.8	0.00
Two or three lobulations	99.2	19.4	19.7	99.2	32.7	0.05
Ellipsoid	97.6	51.2	28.6	99.1	59.2	0.05
Thin capsule	95.2	76.0	44.2	98.8	79.2	0.07

Numbers reflect percentages.
PPV = positive predictive value; NPV = negative predictive value; OR = odds ratio.
Classification of a solid nodule as benign required lack of malignant characteristics, plus hyperechogenicity or a thin echogenic capsule plus ellipsoid shape, or a thin echogenic capsule plus two or three gentle lobulations.
Adapted from Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 1995;196:123–34; with permission.

Table 8
Inter- and intraobserver variability in evaluation of sonography of solid breast masses

	Interobserver	Reproducibility	Intraobserver	Reproducibility
Echogenic pseudocapsule	0.09	Slight	0.63	Substantial
Echogenicity	0.40	Fair	0.69	Substantial
Margin	0.43	Moderate	0.62	Substantial
Echotexture	0.44	Moderate	0.24	Fair
Acoustic transmission	0.55	Moderate	0.63	Substantial
Shape	0.80	Substantial	0.79	Substantial
Final assessment	0.51	Moderate	0.66	Substantial

Data reflect the κ statistic, with <0.2 indicating slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.0, almost perfect agreement.
Adapted from Baker JA, Kornguth PJ, Soo MS, Walsh R, Mengoni P. Sonography of solid breast lesions: observer variability of lesion description and assessment. AJR Am J Roentgenol 1999;172:1621–5; with permission.

mass shape. Variability in descriptions contributed to interobserver and intraobserver inconsistency in assessing the likelihood of malignancy. It is likely the interobserver variability would be even higher if real-time imaging were incorporated into the analysis. Additional work will be necessary to evaluate the interobserver and intraobserver variability in the finalized version of the ACR breast ultrasound lexicon.

Toward a lexicon for breast MRI

Lexicon development

Magnetic resonance imaging of the breast has high sensitivity in the detection of breast cancer, reported as up to 100% in some series, but has lower specificity, ranging from 37% to 97% [33]. Parenchymal breast MRI is also an expensive exa-

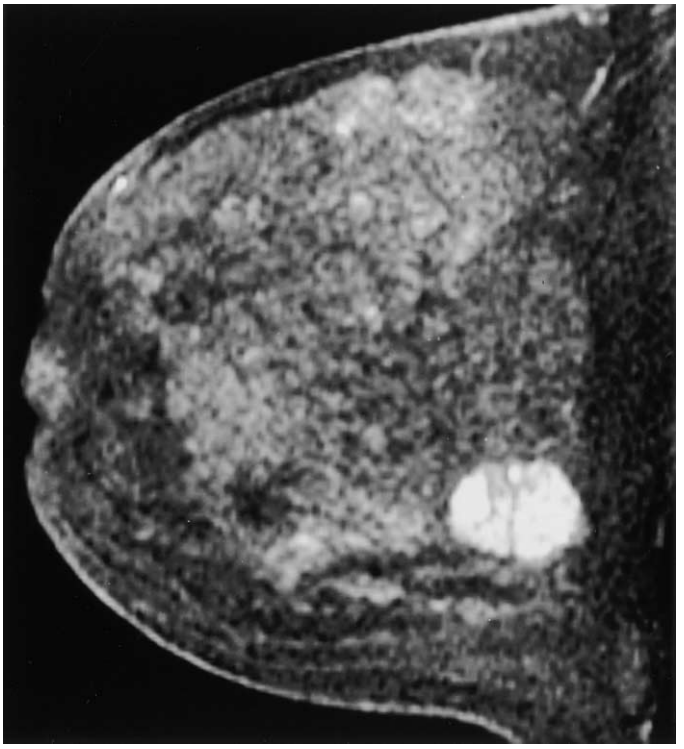

Fig. 10. MRI of fibroadenoma. Sagittal, T1-weighted, contrast-enhanced image shows a lobulated enhancing mass with non-enhancing internal septations.

mination that requires the injection of intravenous contrast material. Published work supports the usefulness of MRI in specific scenarios, such as identification of occult carcinoma, problem-solving, local staging of breast cancer (including skin or pectoral muscle involvement), and (potentially) high-risk screening [33,34]. Progress in breast MRI has been limited by lack of standardization in image acquisition and image interpretation, with some methods focusing on morphology (spatial resolution) and others stressing kinetics (temporal resolution).

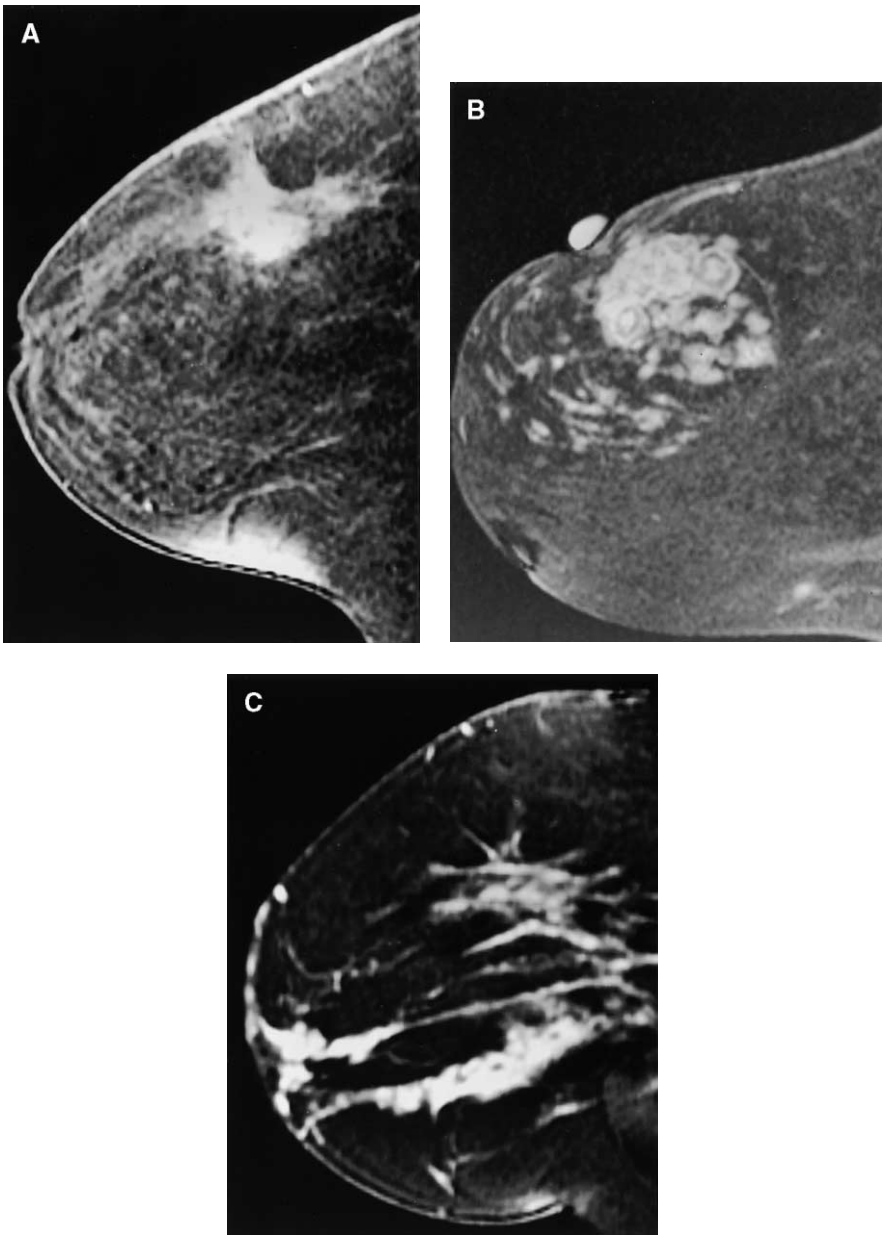

Fig. 11. MRI patterns of breast cancers in sagittal, T1-weighted, contrast-enhanced images. (A) Spiculated, irregular enhanced mass in superior breast; biopsy yielded infiltrating ductal carcinoma and ductal carcinoma in situ (DCIS). Note suboptimal fat suppression inferiorly. (B) Lobulated mass with heterogeneous and rim enhancement; histologic analysis yielded infiltrating ductal carcinoma and DCIS. (C) Extensive linear and segmental clumped enhancement; biopsy yielded DCIS.

In 1997, Nunes et al [35] analyzed the diagnostic accuracy of specific architectural (morphologic) features identified during breast MRI in 93 women. Architectural features that were highly predictive of benign disease included smooth or lobulated borders (97% to 100%), the absence of mass enhancement (100%), and enhancement that was less than the enhancement of surrounding fibroglandular tissue (93% to 100%). Nonenhancing internal septations, present in 9 of 14 (64%) fibroadenomas in a subsequent study, were specific for the diagnosis of fibroadenoma and correlated with collagenous bands at histologic analysis (Fig. 10). Architectural features that were highly predictive of carcinoma included spiculated borders (76–88%) and peripheral rim enhancement in the presence of central lesion enhancement (79–92%) (Fig. 11).

In 1999, Kuhl et al [36] assessed the relevance of signal-intensity time-course analysis (kinetics) for the differential diagnosis of enhancing lesions in breast MRI in a study of 266 breast lesions, of which 101 (40.0%) were malignant. They classified enhancement curves as type 1, steady; type 2, plateau; or type 3, washout (Fig. 12). A washout pattern was significantly more frequently observed in cancers than in benign lesions (Table 9). The diagnostic indices for time signal intensity curves were sensitivity, 91%; specificity, 83%; and diagnostic accuracy, 86%. There was almost perfect interobserver agreement in categorizing the shape of the time signal intensity curve, with $\kappa = 0.85$. The shape of the time signal intensity curve was a more useful predictor of malignancy than the rate of enhancement (Table 9).

Supported by the Office of Women's Health and the ACR, The International Working Group on Breast MRI Imaging is developing a lexicon of terms for breast MRI reporting, the first version of which was published in 1999 [37]. Schnall and Ikeda [37] suggested that MRI reports include descriptions of clinical abnormalities, previous biopsies, hormonal status,

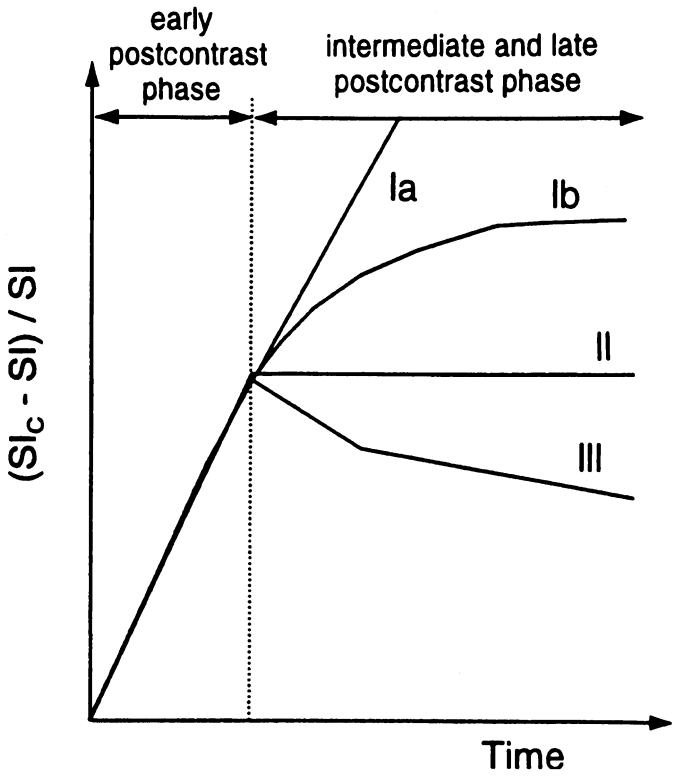

Fig. 12. Schematic drawing of time-signal intensity curve types. Type 1 (persistent or steady) corresponds to a straight (1a) or curved (1b) line; enhancement continues over the entire dynamic study. Type II is a plateau curve with a sharp bend after the initial upstroke. Type III is a washout time course. SI_c = signal intensity after contrast enhancement; SI = signal intensity before contrast injection. (*Data from* Kuhl CK, Mielcareck P, Klaschik S, et al. Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions? Radiology 1999;211:101–110; with permission.)

Table 9
Breast MRI: time signal intensity curves as predictors of malignancy

	Cancers (n = 101)	Benign lesions (n = 165)
Time signal intensity curve		
Type I (steady)	8.9	83.0
Type II (plateau)	33.6	11.5
Type III (washout)	57.4	5.5
Enhancement rate		
Slow	9.0	36.9
Intermediate	25.7	28.5
Fast	65.3	34.5

Numbers reflect the proportion of cancers or benign lesions that had the kinetic features shown. Enhancement rate was defined as the signal intensity increase on the first postcontrast image, with slow being an increase less than or equal to 60%, intermediate being an increase of more than 60% and less than or equal to 80%, and fast being an increase of more than 80%.
Adapted from Kuhl CK, Mielcareck P, Klaschik S, et al. Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions? Radiology 1999;211:101–110; with permission.

and comparison with prior studies. Technical factors should be stated, including the location of markers and significance, magnet field strength, use of a dedicated breast coil, contrast media, pulse sequence, anatomy (including slice thickness and scan orientation and plane), and post-processing techniques. Findings described should include mention of artifacts that affect interpretation, breast composition, implants, and presence or absence of abnormal enhancement, with specific descriptors defined for focal enhancement, kinetics, summary impression, and recommendations. Descriptive terms for breast MRI were elegantly illustrated by Morris [38].

Descriptive terms for breast MRI

Focus/foci
Mass margin
- Smooth
- Irregular
- Spiculated

Mass shape
- Oval
- Round
- Lobulated
- Irregular

Mass enhancement
- Homogeneous
- Heterogeneous
- Rim
- Dark internal septations
- Enhancing internal septations
- Central enhancement

Non-mass enhancement
- Linear (smooth, irregular, or clumped)
- Segmental
- Regional
- Multiple regions
- Diffuse

Non-mass enhancement descriptors for all other types
- Homogeneous
- Heterogeneous
- Stippled/punctate
- Clumped
- Septal/dendritic

Symmetric versus asymmetric for bilateral studies

Adapted from Morris EA. Illustrated breast MR lexicon. In: Miller WT, Bert WA, editors. Seminars in roentgenology. Breast imaging. Vol. 36. Philadelphia: WB Saunders; 2001. p. 238–49; with permission.

Although limited data validate the assignment of final assessment categories based on MRI findings, guidelines were suggested by Kuhl et al [39] in an investigation of breast MRI for high-risk screening. In that study, BI-RADS category 1 was assigned to lesions without any contrast material enhancement. BI-RADS category 2 was assigned to lesions in which enhancement was detected but was classified as benign (focal masses with well-circumscribed morphology, internal septations but otherwise homogeneous enhancement, steady time-signal intensity course, and centrifugal progression of enhancement; or non-mass-related gradual enhancement). BI-RADS category 3 was assigned to lesions compatible with ''unidentified bright objects'' or UBOs (spontaneous, hormone-induced enhancement) and in lesions with presumably benign masses that lacked some of the BI-RADS category 2 features.

BI-RADS category 4 was assigned to lesions with a washout time course, irrespective or morphology, or lesions with suspicious morphology, irrespective of

kinetics. Morphology was suspect if there was spiculated or irregular lesion configuration, heterogeneous internal architecture (particularly rim enhancement), and asymmetric segmental or linear enhancement (see Fig. 11). BI-RADS category 5 was attributed to lesions in which morphologic and architectural features were suggestive of malignancy. Further work is needed to validate this approach.

Potential usefulness of the breast MRI lexicon

Preliminary work supports the usefulness of a BI-RADS lexicon for MRI-detected lesions. Kim et al [40] described the magnetic resonance appearance of 72 focally enhancing infiltrating breast carcinomas. They reported that mass margins were spiculated in 34 (47%), indistinct in 22 (31%), circumscribed in 15 (21%), and obscured in 1 (1%). Mass shape was irregular in 41 (57%), lobular in 16 (22%), round in 10 (14%), and oval in 5 (7%). Enhancement pattern was heterogeneous in 43 (60%), homogeneous in 15 (21%), and rim in 14 (19%). BI-RADS final impression was 3 in 3 (4%), 4 in 26 (36%), and 5 in 43 (60%). There was moderate interobserver agreement for mass margins ($\kappa = 0.46$), mass shape ($\kappa = 0.41$), and enhancement pattern ($\kappa = 0.56$).

Siegmann et al [41] reviewed MRI and histologic findings in 70 exclusively MRI-detected lesions that were prospectively classified as BI-RADS analogous class 3 (probably benign), class 4 (suspicious), or class 5 (highly suggestive of malignancy). The frequency of carcinoma was 0% (0 of 4) for class 3, 23.7% (14 of 59) for class 4, and 85.7% (6 of 7) for class 5, comparable to the frequency of carcinoma for analogous classes in studies of the BI-RADS lexicon for mammography [42–44]. Few details are given regarding criteria for assigning different final assessment categories; this should be clarified in future work.

Precise definition of terms facilitates studies into PPV of specific MRI features. Morakkabati et al [45] reported a pattern of segmental or ductal enhancement in 19 (3.8%) of 500 consecutive patients who underwent dynamic breast MRI. Segmental enhancement occurred in 14 of 19 patients, 10 of whom had DCIS and 4 of whom had fibrocystic change. Ductal enhancement was seen in 5 of 19 patients, 1 of whom had DCIS and 4 of whom had benign findings (1 papilloma and 3 fibrocystic change). The PPV of segmental or ductal enhancement was 58% (11 of 19), and the specificity of this criterion was 98% (481 of 489). The authors concluded that ductal or segmental enhancement was an infrequent finding on breast MRI but that it had high PPV for malignancy.

The breast MRI lexicon is a work in evolution. Standardization of technique would help in the development of the breast MRI lexicon. Further research into the positive and negative predictive values of specific MRI features will be of great value in the complex business of interpreting breast MR images and would allow more women to benefit from the use of breast MRI in the detection and local staging of breast cancer.

Summary and future directions

The BI-RADS lexicon was created to standardize mammographic reporting, thereby enabling better communication, improving clarity in reporting, and facilitating research. The lexicon has enabled studies that have better defined the positive predictive value of specific mammographic features and has contributed to progress in computer-aided diagnosis. In spite of the lexicon's goal of standardization, considerable interobserver and intraobserver variability in mammographic interpretation persists. Further work is necessary to refine the lexicon, to assess training techniques for lexicon use, and to further develop and validate lexicons for breast sonography, breast MRI, and other new imaging modalities as they become available.

References

[1] American College of Radiology. Breast imaging reporting and data system (BI-RADS). Reston, VA: American College of Radiology; 1995.
[2] American College of Radiology. Illustrated breast imaging reporting and data system (BI-RADS™). Reston, VA: American College of Radiology; 1998.
[3] D'Orsi CJ, Kopans DB. Mammographic feature analysis. Semin Roentgenol 1993;28:204–30.
[4] Lacquement MA, Mitchell D, Hollingsworth AB. Positive predictive value of the Breast Imaging Reporting and Data System. J Am Coll Surg 1999;189:34–40.
[5] Liberman L, Abramson AF, Squires FB, Glassman J, Morris EA, Dershaw DD. The Breast Imaging Reporting and Data System: positive predictive value of mammographic features and final assessment categories. AJR Am J Roentgenol 1998;171:35–40.
[6] Orel SG, Kay N, Reynolds C, Sullivan DC. BI-RADS categorization as a predictor of malignancy. Radiology 1999;211:845–50.
[7] Caplan LS, Blackman D, Nadel M, Monticciolo DL. Coding mammograms using the classification "probably benign" finding short interval follow-up suggested. AJR Am J Roentgenol. 1999;172:339–42.
[8] Sickles EA. Periodic mammographic follow-up of

probably benign lesions: results of 3,184 consecutive cases. Radiology 1991;179:463–8.
[9] Sickles EA. Nonpalpable, circumscribed, noncalcified solid breast masses: likelihood of malignancy based on lesion size and age of patient. Radiology 1994;192: 439–42.
[10] Varas X, Leborgne F, Leborgne JH. Nonpalpable, probably benign lesions: role of follow-up mammography. Radiology 1992;184:409–14.
[11] Sickles EA. Probably benign breast lesions: when should follow-up be recommended and what is the optimal follow-up protocol? Radiology 1999;213:11–4.
[12] Rubin E. Six-month follow-up: an alternative view. Radiology 1999;213:15–8.
[13] Saftlas AF, Wolfe JN, Hoover RN, Brinton LA, Schairer C, Salane M, et al. Mammographic parenchymal patterns as indicators of breast cancer risk. Am J Epidemiol 1989;129:518–26.
[14] Whitehead J, Carlile T, Kopecky KJ, Thompson DJ, Gilbert FI Jr, Present AJ, et al. The relationship between Wolfe's classification of mammograms, accepted breast cancer risk factors, and the incidence of breast cancer. Am J Epidemiol 1985;122:994–1006.
[15] Wolfe JN. Breast patterns as an index of risk for developing breast cancer. AJR Am J Roentgenol 1976;126: 1130–9.
[16] Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92:1081–7.
[17] Satija S, Moore RH, Michaelson JS, Halpern E, Slanetz PJ, Kopans DB. Breast tissue patterns as a risk factor for developing cancer [abstract]. Radiology 2000; 217(P):446.
[18] Baker JA, Kornguth PJ, Lo JY, Williford ME, Floyd CE Jr. Breast cancer: prediction with artificial neural network based on BI-RADS standardized lexicon. Radiology 1995;196:817–22.
[19] Baker JA, Kornguth PJ, Lo JY, Floyd CE Jr. Artificial neural network: improving the quality of breast biopsy recommendations. Radiology 1996;198:131–5.
[20] Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in radiologists' interpretations of mammograms. N Engl J Med 1994;331:1493–9.
[21] Beam CA, Layde PM, Sullivan DC. Variability in the interpretation of screening mammograms by US radiologists: findings from a national sample. Arch Intern Med 1996;156:209–13.
[22] D'Orsi CJ, Swets JA. Variability in the interpretation of mammograms [letter]. N Engl J Med 1995;332:1172.
[23] Kopans DB. The accuracy of mammographic interpretation. N Engl J Med 1994;331:1521–2.
[24] Baker JA, Kornguth PJ, Floyd CE Jr. Breast Imaging Reporting and Data System standardized mammography lexicon: observer variability in lesion description. AJR Am J Roentgenol 1996;166:773–8.
[25] Kerlikowske K, Grady D, Barclay J, Fraakel SD, Dminsky SH, Sickles EA, et al. Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System. J Natl Cancer Inst 1998;90:1801–9.
[26] Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast Imaging Reporing and Data System: inter- and intraobserver variability in feature analysis and final assessment. AJR Am J Roentgenol 2000;174: 1769–77.
[27] Berg WA, D'Orsi CJ, Jackson VP, Bassett LW, Beam CA, Crewson PE. Effect of training in BI-RADStm on reader agreement in feature analysis in mammography [abstract]. Radiology 2000;217(P):571.
[28] Vitiello SM, Philpotts LE, Tocino I, Horvath LJ, Lee CH. Understanding of BI-RADS among referring clinicians: do they get it [abstract]? AJR Am J Roentgenol 2000;174(Suppl):64.
[29] Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 1995;196:123–34.
[30] Berg WA, Fajardo LL, Pisano ED, Gatsonis C, McNeil BJ. Correlation of sonographic findings with risk of malignancy: the RDOGV experience [abstract]. Radiology 1999;213(P):107.
[31] Mendelson EB, Berg WA, Merrit CRB. Toward a standardized breast ultrasound lexicon, BI-RADS: ultrasound. In: Miller WT, Berg WA, editors. Seminars in roentgenology, vol 36: breast imaging. Philadelphia, PA: WB Saunders; 2001. p. 217–25.
[32] Baker JA, Kornguth PJ, Soo MS, Walsh R, Mengoni P. Sonography of solid breast lesions: observer variability of lesion description and assessment. AJR Am J Roentgenol. 1999;172:1621–5.
[33] Orel SG, Schnall MD. MR imaging of the breast for the detection, diagnosis, and staging of breast cancer. Radiology 2001;220:13–30.
[34] Morris EA. Review of breast MRI: indications and limitations. In: Miller WT, Berg WA, editors. Seminars in roentgenology, vol. 36: breast imaging. Philadelphia, PA: WB Saunders; 2001. p. 226–37.
[35] Nunes LW, Schnall MD, Orel SG, Hachman MG, Langlotz CP, Reynolds CA, et al. Breast MR imaging: interpretation model. Radiology 1997;202:833–41.
[36] Kuhl CK, Mielcareck P, Klaschik S, Leutner C, Wardelmann E, Gieseke J, et al. Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions? Radiology 1999;211:101–10.
[37] Schnall MD, Ikeda DM. Lesion diagnosis working group report. J Magn Reson Imaging 1999;10:982–90.
[38] Morris EA. Illustrated breast MR lexicon. In: Miller WT, Berg WA, editors. Seminars in roentgenology, vol. 36: breast imaging. Philadelphia, PA: WB Saunders; 2001. p. 238–49.
[39] Kuhl CK, Schmutzler RK, Leutner CC, Kempe A, Wardelmann E, Hocke A, 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;215:267–79.

[40] Kim SJ, Morris EA, Liberman L, Ballon DJ, LaTrenta LR, Hadar O, et al. Magnetic resonance appearance of focal infiltrating carcinomas [abstract]. Radiology 1999;213(P):454.

[41] Siegmann KC, Fersis N, Ruck P, Claussen CD, Mueller-Schimpfle MP. Positive predictive value of MR characteristics and BI-RADS-analogous categories of exclusively MR detected breast lesions [abstract]. Radiology 2000;217(P):524.

[42] Lacquement MA, Mitchell D, Hollingsworth AB. Positive predictive value of the Breast Imaging Reporting and Data System. J Am Coll Surg 1999;189:34–40.

[43] Liberman L, Abramson AF, Squires FB, Glassman J, Morris EA, Dershaw DD. The Breast Imaging Reporting and Data System: positive predictive value of mammographic features and final assessment categories. AJR Am J Roentgenol. 1998;171:35–40.

[44] Orel SG, Kay N, Reynolds C, Sullivan DC. BI-RADS categorization as a predictor of malignancy. Radiology 1999;211:845–50.

[45] Morakkabati N, Schmiedel A, Leutner C, Kuhl CK. Diagnostic usefulness of ductal or segmental enhancement in dynamic breast MR imaging [abstract]. Radiology 2000;217(P):526–27.