BI-RADS Lexicon for US and Mammography: Interobserver Variability and Positive Predictive Value[1]

Elizabeth Lazarus, MD
Martha B. Mainiero, MD
Barbara Schepps, MD
Susan L. Koelliker, MD
Linda S. Livingston, MD

Purpose: To retrospectively evaluate interobserver variability between breast radiologists by using terminology of the fourth edition of the Breast Imaging Reporting and Data System (BI-RADS) to categorize lesions on mammograms and sonograms and to retrospectively determine the positive predictive value (PPV) of BI-RADS categories 4a, 4b, and 4c.

Materials and Methods: Institutional review board approval was obtained; informed consent was not required. This study was HIPAA compliant. Ninety-four consecutive lesions in 91 women who underwent image-guided biopsy comprised 59 masses, 32 calcifications, and three masses with calcification. Five radiologists retrospectively reviewed these lesions. Each observer described each lesion with BI-RADS terminology and assigned a final BI-RADS category. Interobserver variability was assessed with the Cohen κ statistic. A pathologic diagnosis was available for all 94 lesions; 30 (32%) were malignant and 64 (68%) were benign. Pathologic analysis of benign lesions was performed on tissue obtained with image-guided core-needle biopsy. In cases referred for excisional biopsy after needle biopsy because of atypia or discordance, final surgical pathologic analysis was used for correlation with imaging findings. PPV for category 4 or 5 lesions was determined for all readers combined.

Results: For ultrasonographic (US) descriptors, substantial agreement was obtained for lesion orientation, shape, and boundary (κ = 0.61, 0.66, and 0.69, respectively). Moderate agreement was obtained for lesion margin and posterior acoustic features (κ = 0.40 for both). Fair agreement was obtained for lesion echo pattern (κ = 0.29). For mammographic descriptors, moderate agreement was obtained for mass shape, mass margin, and calcification distribution (κ = 0.48, 0.48, and 0.50, respectively). Fair agreement was obtained for calcification description (κ = 0.32). Slight agreement was obtained for mass density (κ = 0.18). Fair agreement was obtained for final assessment category (κ = 0.28). PPVs of BI-RADS category 4 and 5 assignments were as follows: category 4a, six (6%) of 102; category 4b, 17 (15%) of 110; category 4c, 48 (53%) of 91; and category 5, 71 (91%) of 78.

Conclusion: Interobserver agreement with the new BI-RADS terminology is good and validates the US lexicon. Subcategories 4a, 4b, and 4c are useful in predicting the likelihood of malignancy.

© RSNA, 2006

The Breast Imaging Reporting and Data System (BI-RADS) developed by the American College of Radiology (1) has standardized the assessment and reporting of breast lesions identified on mammograms. Both lesion description and management recommendations have become more consistent with use of BI-RADS. By making auditing easier, BI-RADS categorization of reports has also facilitated quality assurance practices (2,3).

Breast ultrasonography (US) has proved to be useful in the evaluation of masses detected with mammography or clinical examination, as US is used to distinguish cystic lesions from solid lesions and to further differentiate benign solid masses from malignant solid masses (4–6). However, standard terminology for describing lesions on breast sonograms has been lacking. Use of BI-RADS for breast US should standardize the reporting and classification of lesions detected on sonograms, thereby improving the utility of US in the work-up of breast masses. By identifying lesion descriptors that emphasize the distinctions between benign and malignant lesions on breast sonograms, the BI-RADS lexicon for US clarifies the indication for biopsy of particular lesions.

One other major addition in the fourth edition of BI-RADS involves the subcategorization of category 4 lesions. While classification of a lesion as BI-RADS category 4 indicates that a lesion has been recommended for biopsy, it provides no frame of reference for either the referring physician or the patient as to the prebiopsy risk for malignancy. Dividing category 4 lesions into those with a small (category 4a), moderate (category 4b), or substantial (category 4c) (1) likelihood of malignancy better informs the physician and patient as to the level of concern regarding the lesion and prepares both the physician and the patient for the likely biopsy findings and the potential need for follow-up.

Our study had two purposes: to retrospectively evaluate interobserver variability between breast radiologists who used the new BI-RADS terminology to characterize lesions identified on both mammograms and sonograms and to retrospectively determine the positive predictive value of the new BI-RADS categories (4a,

4b, and 4c) as they are used by radiologists performing breast imaging.

Materials and Methods

Institutional review board approval was obtained, and the informed consent requirement was waived; this study was compliant with the Health Insurance Portability and Accountability Act.

Patient Lesions and Interpretation

The authors (all radiologists with subspecialty expertise in breast imaging) retrospectively evaluated 94 consecutive lesions in 91 women (mean age, 55 years; age range, 28–85 years) who underwent image-guided biopsy between August 1, 2002 and October 4, 2002. Four of the radiologists (E.L., M.B.M., L.S.L., and S.L.K.) underwent fellowship training in breast imaging and have practiced as faculty in an academic breast imaging section for 2–10 years. The other radiologist (B.S.) has been interpreting breast images for more than 30 years in her position as chief of the same breast imaging section. All five radiologists practice within the same group, and all met the standards of the Mammography Quality Standards Act as qualified interpreting physicians. The 94 lesions comprised 59 masses, 32 calcifications, and three masses with calcification; 52 lesions were evaluated by reviewing both mammograms and sonograms, 32 were evaluated by reviewing mammograms alone, and 10 were evaluated by reviewing sonograms alone. The lesions were evaluated on original mammograms and sonograms marked to indicate lesions, without the benefit of prior mammograms and sonograms for comparison. The lesions were marked so that all observers would examine the same biopsy-proved lesion, as some images showed more than one finding. Mammographic magnification views were obtained for all calcifications analyzed. All 62 masses were evaluated with US; 52 (84%) masses were evaluated with both mammography and US. Of the 94 lesions, 30 (32%) were malignant and 64 (68%) were benign.

Pathologic diagnosis was available for all lesions. Pathologic findings in benign lesions were evaluated with results from image-guided core-needle biopsy; a 14-gauge needle was used for US-guided bi-

opsy, and a 9-gauge needle was used for vacuum-assisted stereotactic biopsy. In cases referred for excisional biopsy after needle biopsy because of atypia, positivity for malignancy, or discordance, final surgical pathologic analysis was used for correlation with imaging findings.

Each observer described each lesion by using the terminology of the fourth edition of the BI-RADS lexicon and assigned a final BI-RADS category, including the new subcategories of BI-RADS category 4. Each observer was provided a sheet containing the BI-RADS categories and descriptors for lesions seen on both mammograms and sonograms and instructed to select the most appropriate descriptors for each lesion.

While the evaluating radiologists were familiar with the "Guidance Chapter" (1) criteria for BI-RADS subcategorization, no formal training in the newest lexicon was provided. Thus, the criteria used by each evaluating radiologist were subjective and based on prior knowledge and experience. Table 1 lists the BI-RADS terminology used in this study.

Statistical Analysis

The Cohen κ statistic was used to assess interreader agreement for all descriptor variables. The guidelines of Landis and Koch were followed in interpreting κ values: 0.00–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.80–1.00, almost perfect agreement (7). All κ statistics were calculated with statistical software (Stata,

Radiology

version 8; Stata, College Station, Tex). The positive predictive value for category 4 and 5 lesions was determined by using data from the assessments of all readers combined.

To determine if the results were skewed by any readers, we assessed the reliability between pairs of readers. Data from the final BI-RADS categorization of each lesion were used to determine the Pearson product-moment correlation coefficients between each pair of readers. These statistics were calculated with the SAS system (SAS Institute, Cary, NC). Intraclass correlation was also determined for final BI-RADS categorization by using *(a)* all results from each reader and *(b)* the mean of the results from each reader (Appendix).

Results

Mammographic Assessment of Masses

In describing the mass shape, overall agreement was moderate (κ = 0.48). Substantial agreement was seen when the mass shape was characterized as irregular (κ = 0.68), moderate agreement was seen when the mass shape was characterized as oval (κ = 0.46), and fair agreement was seen when the mass shape was characterized as lobular (κ = 0.24) or round (κ = 0.31) (Table 2).

In assessing the margins of the mass, overall agreement was moderate (κ = 0.48). The greatest agreement was seen when the margins were characterized as circumscribed (κ = 0.60) or spiculated (κ = 0.69). Poor agreement was seen when the margins were characterized as microlobulated (κ = 0.08), while fair agreement was seen when margins were characterized as indistinct (κ = 0.38) or obscured (κ = 0.27).

Overall agreement for mass density was slight (κ = 0.18). Agreement was slight when mass density was described as fat containing (κ = 0.11), equal to breast tissue (κ = 0.15), or high (κ = 0.20).

Mammographic Assessment of Calcifications

Agreement was nearly perfect when assessing the presence of calcifications (κ = 0.94) (Table 2). The five observers dem-

onstrated overall fair agreement when they described the calcifications (κ = 0.32). Use of the terms *amorphous* and *fine branching* resulted in moderate agreement (κ = 0.45 and 0.49, respectively). Agreement was fair for use of the terms *coarse heterogeneous* (κ = 0.27), *pleomorphic* (κ = 0.21), and *vascular* (κ = 0.24). Use of the terms *coarse* (*n* = 3), *dystrophic* (*n* = 1), *milk of calcium* (*n* = 2), and *punctuate* (*n* = 2) was uncommon in this series of lesions and led to low κ values (κ = −0.01 for these terms). Figure 1 shows an image for which reviewers agreed on the number and distribution of calcifications but disagreed on the description of calcifications.

Moderate agreement was achieved in the evaluation of the distribution (κ = 0.50) and number (κ = 0.48) of calcifications.

In evaluating the presence of architectural distortion, agreement was fair (κ = 0.26). Agreement between readers for the presence of associated findings and special cases could not be assessed secondarily because the readers found few cases in which associated findings or special cases were present in the lesions.

Sonographic Assessment

For sonographic descriptors, substantial agreement was obtained for assessment of lesion orientation (κ = 0.61) that was described as parallel or not parallel (Table 3).

Substantial agreement was also found for evaluation of lesion shape (κ = 0.66). The greatest agreement was achieved when lesion shape was described as irregular (κ = 0.70) or oval (κ = 0.71). The term *round* was used infrequently, and less agreement was found with its use (κ = 0.29).

Evaluation of the lesion boundary, which was described as abrupt or having an echogenic halo, yielded similarly substantial agreement (κ = 0.69).

Fair agreement was achieved for evaluation of the lesion margin (κ = 0.40). Excellent agreement was seen with lesions that were considered circumscribed (κ = 0.71). Fair agreement was seen with lesions that were considered angular (κ = 0.22), indistinct (κ = 0.22), microlobulated (κ = 0.25), or spiculated (κ = 0.26).

The many terms available to describe the US echo pattern yielded fair agreement between observers (κ = 0.29). The terms *complex* (κ = 0.40) and *hypoechoic* (κ = 0.29) were used most commonly and demonstrated the most agreement in their use. The terms *anechoic* (κ = −0.01) and *isoechoic* (κ = 0.05) were used rarely, which likely accounted for poor agreement in their use. The term *hyperechoic* also was used rarely and demonstrated only slight agreement (κ = 0.16).

Overall fair agreement was achieved in describing posterior acoustic features (κ = 0.40). The greatest agreement (κ = 0.66) was achieved when lesions were described as having posterior acoustic shadowing. When lesions were described as having either posterior enhancement or no change in the echo pattern, agreement was fair (κ = 0.39 and κ = 0.31, respectively). Lesions were rarely described as having combined posterior enhancement and shadowing, and agreement between observers who used this description was poor (κ = 0.09). Figure 2 shows a lesion with good agreement for lesion orientation, shape, boundary, and echogenicity but disagreement for assessment of the lesion margin and posterior acoustic features.

No κ statistics could be calculated for assessment of alterations in the surrounding tissue, presence of calcifications, lesions determined to be special cases, or lesion vascularity because the observers believed these findings were present in the lesions only on rare occasions.

Final Assessment Category

Fair agreement was achieved for the final assessment category (κ = 0.28) with all final categories (2, 3, 4a, 4b, 4c, and 5) (Table 4). The greatest agreement was found with lesions categorized as highly suspicious for malignancy (category 5) (κ = 0.56). Although few lesions were rated as benign (category 2), agreement with this category was fair (κ = 0.27). Fair agreement was obtained for category 3 and 4c lesions (κ = 0.32 and 0.26, respectively); however, there was poor agreement between observers for category 4a and 4b lesions (κ = 0.14 and 0.16, respectively). When the categories were grouped in terms of whether biopsy

was required (categories 2 and 3 and categories 4a, 4b, and 5), moderate agreement was obtained (κ = 0.45). Figure 3 shows a classic malignancy on a sonogram, with all reviewers classifying the malignancy as BI-RADS category 5.

The results of the Pearson product-moment correlation demonstrate fair agreement between all pairs of readers (Table 5). This indicates that the results of no one reader skewed the analysis. Intraclass correlation with all final BI-RADS assessments showed moderate reliability of 0.676. Use of each reader's mean value of BI-RADS assessment produced improved reliability, with an intraclass correlation of 0.912.

Positive Predictive Value

Pathologic findings in all lesions were available for review. Forty-eight (51%) of the lesions were diagnosed with core-needle biopsy only. In 46 (49%) lesions, pathologic analysis was based on the findings of excisional biopsy. With the results of all readers combined, positive predictive values of lesions categorized as BI-RADS 4 or 5 were as follows: category 4a, six (6%) of 102; category 4b, 17 (15%) of 110; category 4c, 48 (53%) of 91; and category 5, 71 (91%) of 78.

Discussion

Breast US is now considered integral in the assessment of a mass detected with clinical examination and mammography (5). Stavros et al (5) identified individual sonographic characteristics—including shape, margin, lesion boundary, lesion orientation, and echotexture—that indicate whether a lesion is malignant, benign, or indeterminate. With the use of these features, observers have successfully used US criteria to predict malignancy in a breast mass (4). Until recently, however, there has been a lack of uniformity in descriptive terms for US lesions, which can result in inconsistent diagnoses. Baker et al (10) examined observer variability in the description and assessment of solid breast masses on sonograms and identified the need to standardize terminology for breast US, which has been provided in the fourth edition of BI-RADS.

Table 1

BI-RADS Fourth Edition Terminology

Mammographic Evaluation		Characteristic
Calcifications		
Description		
	Typically benign	Vascular
		Coarse or popcornlike
		Rodlike
		Round
		Punctate
		Lucent center
		Rim or eggshell
		Milk of calcium
		Suture
		Dystrophic
	Intermediate	Amorphous or indistinct
		Coarse heterogeneous
	Higher probability of malignancy	Pleomorphic
		Fine branching or casting
Distribution		Grouped or clustered
		Linear
		Segmental
		Regional
		Diffuse or scattered
Number		<5
		5–10
		>10
Masses		
Shape		Round
		Oval
		Lobular
		Irregular
Margins		Circumscribed
		Microlobulated
		Indistinct or ill defined
		Spiculated
Density		High
		Equal
		Low
		Fat containing
Architectural Distortion		
Special cases		Intramammary lymph node
		Tubular density or dilated duct
		Global asymmetry
		Focal asymmetry
Associated findings		Skin retraction
		Nipple retraction
		Skin thickening
		Trabecular thickening
		Skin lesion
		Axillary adenopathy

(Table 1 continues)

Radiology

Table 1 (continued)

BI-RADS Fourth Edition Terminology

Mammographic Evaluation	Characteristic
US	
Orientation	Parallel
	Not parallel
Shape	Oval
	Round
	Irregular
Lesion boundary	Abrupt interface
	Echogenic halo
Echo pattern	Anechoic
	Hyperechoic
	Complex
	Hypoechoic
	Isoechoic
Posterior acoustic features	None
	Enhancement
	Shadowing
	Combined pattern
Surrounding tissue	Duct changes
	Cooper ligament changes
	Edema
	Architectural distortion
	Skin thickening
	Skin retraction or irregularity
Calcifications	Macrocalcifications
	Microcalcifications in mass
	Microcalcifications out of mass
Special cases	Clustered microcysts
	Complicated cysts
	Mass in or on skin
	Foreign body
	Intramammary lymph node
	Axillary lymph node
Vascularity	Not present or not assessed
	Present in lesion
	Present adjacent to lesion
	Diffusely increased in surrounding tissue

Table 2

Interobserver Variability in Description of Mammographic Lesions

BI-RADS Descriptor	κ Value
Mass	
Presence	0.84
Shape	0.48
Margins	0.48
Density	0.18
Calcifications	
Presence	0.94
Description	0.32
Distribution	0.50
Number	0.48
Architectural distortion presence	0.26

Our results show a high degree of agreement in describing lesions on sonograms, thus demonstrating the appropriateness of the terms chosen in the newest iteration of BI-RADS. The terminology was familiar to radiologists experienced in breast imaging, and their use was generally concordant.

Agreement for lesion orientation, shape, and boundary on sonograms was slightly better than agreement for lesion margin on sonograms and mammograms and lesion shape on mammograms, likely because of the greater number of choices for describing the second set of qualifiers. The lower rate of agreement for lesion echo pattern on sonograms and mass density on mammograms suggests that observers had difficulty in making these categorizations. However, results from Stavros et al (5) and Baker et al (10) indicate that these qualifiers are not very useful in the differentiation of benign and malignant masses.

Agreement for calcification description on mammograms was somewhat lower in our study than in prior studies (11,12). The agreement may be lower because of bias introduced by the method chosen to select the cases. To have a pathologic correlation available for assessment of positive predictive value, we chose to evaluate lesions referred for biopsy in this study. As a consequence, there was rare use of typical benign descriptors, which may have better observer agreement.

The fair agreement for overall BI-RADS category reported in our study (κ = 0.28) was not much different from that reported by Berg et al (11) (κ = 0.37) in a prior study in which only mammography was used. The lack of agreement between the κ value obtained in our study and that obtained in the study of Berg et al (11) can be at least partially explained by the greater number of categories offered with inclusion of the 4a, 4b, and 4c subcategories.

Orel et al (13) showed that placement of mammographic lesions into BI-RADS categories is useful for predicting malignancy, with a positive predictive value of 30% for category 4 lesions and 97% for category 5 lesions. We obtained a similar positive predictive value of 91% for category 5 lesions.

While BI-RADS category 5 has always been used to identify lesions that are almost certainly malignant, BI-RADS category 4 historically has comprised a more heterogeneous population of lesions. Our results demonstrate that the optional subcategories of 4a, 4b, and 4c are useful in stratifying the likelihood of malignancy among the large heterogeneous group of category 4 lesions. This stratification is helpful in communicating the level of suspicion to referring physicians and patients, who may choose to use this information in management decisions (ie, which patients to refer to a breast specialist prior to biopsy). In our practice, some referring internists and gynecologists are willing to discuss the results of benign breast biopsies with their patients, but others prefer that patients receive the results of malignant breast biopsies from a breast surgeon who can counsel the patient on further intervention. The medical expenditure of an additional referral may not be necessary if the preprocedural risk assessment is low and can be clearly communicated.

For radiologists who use the subcategories, a medical audit of the positive pre-
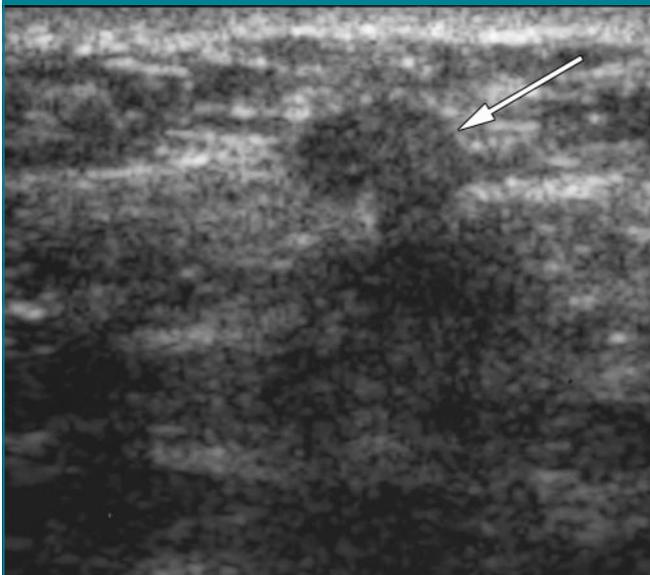
Figure 1
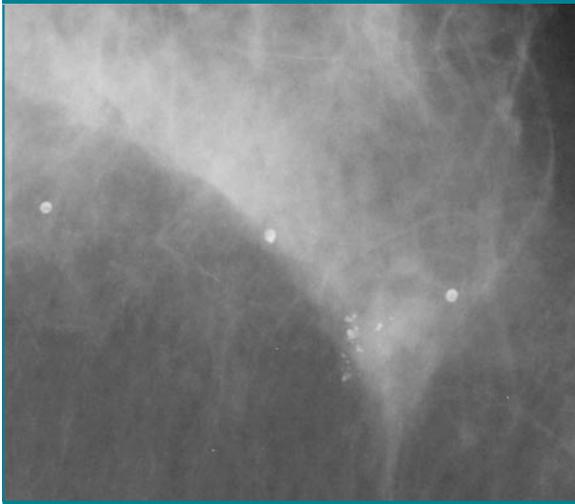
Figure 1: Mammogram shows calcifications in a 70-year-old woman. The calcifications occupy a total area of 6 × 3 mm and were photographically enlarged on this craniocaudal view. All reviewers agreed that the calcifications were grouped in distribution and numbered more than 10; however, the reviewers disagreed concerning calcification description, as they used the terms *coarse* ($n = 1$), *coarse heterogeneous* ($n = 2$), and *pleomorphic* ($n = 2$). Pathologic evaluation after vacuum-assisted stereotactic breast biopsy revealed a degenerating fibroadenoma.

Table 3

Interobserver Variability in Description of US Lesions

BI-RADS Descriptor of Mass	κ Value
Presence	1.00
Orientation	0.61
Shape	0.66
Margin	0.40
Lesion boundary	0.69
Echo pattern	0.29
Posterior acoustic features	0.40

Figures 2, 3

Figure 2: Transverse sonogram in a 46-year-old woman with a solid breast mass (arrow). Four reviewers agreed on lesion orientation (parallel), shape (oval), boundary (abrupt), and echogenicity (hypoechoic). Reviewers did not agree on lesion margin and posterior acoustic features. Pathologic evaluation of core and excisional biopsy samples revealed infiltrating ductal carcinoma.
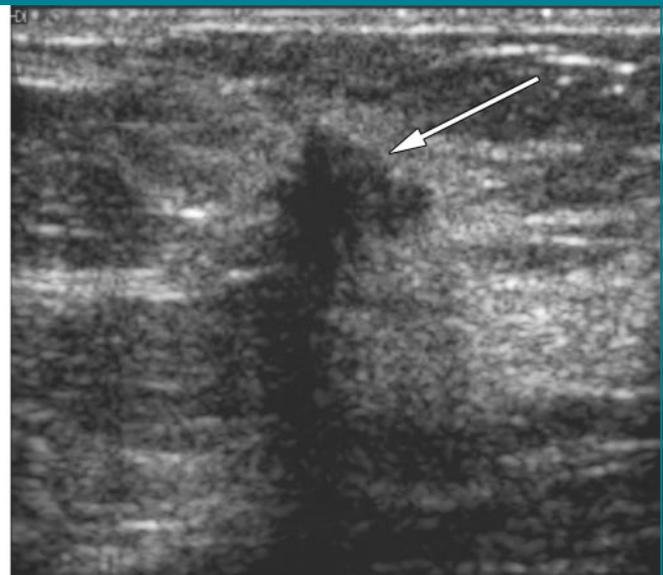
Figure 3: Transverse sonogram in a 56-year-old woman with a solid breast mass (arrow). All reviewers categorized this lesion as BI-RADS category 5. Pathologic evaluation of core and excisional biopsy samples revealed infiltrating ductal carcinoma.

Table 4

Interobserver Variability in Assigning Final BI-RADS Assessment Category

BI-RADS Category	κ Value
2	0.27
3	0.32
4a	0.14
4b	0.16
4c	0.26
5	0.56
Combined*	0.28

* BI-RADS categories 2–5.

Table 5

Pearson Product Moment Correlation Coefficients between All Pairs of Readers

Reader	Reader 1	Reader 2	Reader 3	Reader 4
2	0.78	. . .	. . .	. . .
3	0.62	0.61	. . .	. . .
4	0.76	0.70	0.73	. . .
5	0.72	0.61	0.59	0.66

Note.—For all coefficients, $P < .001$.

dictive value of each category can provide additional feedback on interpretive performance.

There were several limitations to our study. Because we used only those lesions that were referred for biopsy, we had few descriptors typically associated with benign disease and few lesions that could be characterized as special cases. The small number of these cases may have decreased interobserver agreement in some areas, as Taplin et al (2) demonstrated that BI-RADS evaluation of negative findings and benign lesions is more consistent than BI-RADS evaluation of abnormalities.

Second, Berg et al (14) demonstrated that even for experienced breast imagers, BI-RADS training results in improved agreement for feature analysis and final assessment. While the guidance chapter in the fourth edition of the BI-RADS breast imaging atlas offers useful examples of lesions appropriate for each subcategory (1), the evaluating radiologists were not specifically asked to review these criteria. We did not provide specific training to the radiologists involved in this study to more accurately represent the usage of the majority of radiologists being introduced to this new version, who would likely have no formal training prior to its implementation. Additionally, this study was based on the performance of experienced breast imaging radiologists. Inconsistencies and errors in using the BI-RADS lexicon and categories among radiologists with different levels of experience may vary and should be studied.

Another limitation is that the cases had been evaluated by a radiologist prior to this study. We used cases that were assessed more than a year before the study was begun to minimize any recollection of a case that may have influenced categorization of the lesion.

In conclusion, the addition of the BI-RADS lexicon for US is helpful and can be used with good agreement among radiologists, even those without specific training in the new terminology. Additionally, use of the new optional subcategories 4a, 4b, and 4c is beneficial in stratifying the likelihood of malignancy in lesions recommended for biopsy.

Appendix

Under conditions similar to ours (where reasonable numeric values may be assigned to ordinal categories and there are more than three such categories), Hripcsak and Heitjan (8) recommended treating ordinal data as continuous data and examining agreement and reliability by using intraclass and product-moment correlation coefficients. Thus, the 10 orthogonal Pearson product-moment correlation coefficients between readers' judgments were presented, and reliability was estimated with intraclass correlation coefficients for interreader reliability and the reliability of the mean of the readers' judgments by using the methods suggested by Shrout and Fleiss (9).

References

1. American College of Radiology. Breast imaging reporting and data system, breast imaging atlas. 4th ed. Reston, Va: American College of Radiology, 2003.
2. Taplin SH, Ichikawa LE, Kerlikowske K, et al. Concordance of breast imaging reporting and data system assessments and management recommendations in screening mammography. Radiology 2002;222:529–535.
3. Geller BM, Barlow WE, Ballard-Barbash R, et al. Use of the American College of Radiology BI-RADS to report on the mammographic evaluation of women with signs and symptoms of breast disease. Radiology 2002;222:536–542.
4. Rahbar G, Sie AC, Hansen GC, et al. Benign versus malignant solid breast masses: US differentiation. Radiology 1999;213:889–894.
5. 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;196:123–134.
6. Skaane P, Engedal K. Analysis of sonographic features in the differentiation of fibroadenoma and invasive ductal carcinoma. AJR Am J Roentgenol 1998;170:109–114.
7. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.
8. Hripcsak G, Heitjan DF. Measuring agreement in medical informatics reliability studies. J Biomed Inform 2002;35:99–110.
9. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420–428.
10. 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;172:1621–1625.
11. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast imaging reporting and data system: inter- and intraobserver variability in feature analysis and final assessment. AJR Am J Roentgenol 2000;174:1769–1777.
12. Baker JA, Kornguth PJ, Floyd CE. Breast imaging reporting and data system standardized mammography lexicon: observer variability in lesion description. AJR Am J Roentgenol 1996;166:773–778.
13. Orel SG, Kay N, Reynolds C, Sullivan DC. BI-RADS categorization as a predictor of malignancy. Radiology 1999;211:845–850.
14. Berg WA, D'Orsi CJ, Jackson VP, et al. Does training in the breast imaging reporting and data system (BI-RADS) improve biopsy recommendations or feature analysis agreement with experienced breast imagers at mammography? Radiology 2002;224:871–880.