Validation of the fifth edition BI-RADS ultrasound lexicon with comparison of fourth and fifth edition diagnostic performance using video clips

Jung Hyun Yoon[1], Min Jung Kim[1], Hye Sun Lee[2], Sung Hun Kim[3], Ji Hyun Youk[4], Sun Hye Jeong[5], You Me Kim[6]

[1]Department of Radiology, Severance Hospital, Research Institute of Radiological Science, Yonsei University College of Medicine, Seoul; [2]Biostastistics Collaboration Unit, Medical Research Center, Yonsei University College of Medicine, Seoul; [3]Department of Radiology, Seoul St. Mary's Hospital, The Catholic University of Korea, Seoul; [4]Department of Radiology, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul; [5]Department of Radiology, Soonchunhyang University Bucheon Hospital, Bucheon; [6]Department of Radiology, Dankook University Hospital, Dankook University College of Medicine, Cheonan, Korea

Purpose: The aim of this study was to evaluate the positive predictive value (PPV) and the diagnostic performance of the ultrasonographic descriptors in the fifth edition of BI-RADS, comparing with the fourth edition using video clips.

Methods: From September 2013 to July 2014, 80 breast masses in 74 women (mean age, 47.5±10.7 years) from five institutions of the Korean Society of Breast Imaging were included. Two radiologists individually reviewed the static and video images and analyzed the images according to the fourth and fifth edition of BI-RADS. The PPV of each descriptor was calculated and diagnostic performances between the fourth and fifth editions were compared.

Results: Of the 80 breast masses, 51 (63.8%) were benign and 29 (36.2%) were malignant. Suspicious ultrasonographic features such as irregular shape, non-parallel orientation, angular or spiculated margins, and combined posterior features showed higher PPV in both editions (all P<0.05). No significant differences were found in the diagnostic performances between the two editions (all P>0.05). The area under the receiver operating characteristics curve was higher in the fourth edition (0.708 to 0.690), without significance (P=0.416).

Conclusion: The fifth edition of the BI-RADS ultrasound lexicon showed comparable performance to the fourth edition and can be useful in the differential diagnosis of breast masses using ultrasonography.

Keywords: Breast neoplasms; Biopsy; Breast Imaging Reporting and Data System (BI-RADS); Ultrasonography

Introduction

The role of breast ultrasonography (US) has rapidly expanded from simply characterizing the internal contents of the mass to differentiating between benign to malignant breast masses and as an adjunctive to mammography. It has even been proposed as a screening modality in young women or women with dense breasts [1–4]. With the widespread use of breast US in everyday practice, the American College of Radiology organized and released the Breast Imaging Reporting and Data System (BI-RADS) lexicon for breast US in 2003 in order to provide standardized lesion characterization and US reporting, facilitate better communication between radiologists and clinicians, and bring further uniformity to recommendations [5]. This lexicon uses descriptors such as shape, orientation, margin, lesion boundary, echo pattern, posterior features, and calcifications in lesion description for breast masses detected on US [5], which has been proven both effective and feasible in breast mass characterization by many studies [2,6,7].

In 2013, the fifth edition of BI-RADS [8] was released, containing several changes in the US section. "Heterogeneous" echogenicity has been added as an echo pattern, defined as a mixture of echogenic patterns within a solid mass [8]. The "lesion boundary" category has been eliminated, and "intraductal calcifications" has been added in calcification descriptors, reflecting the usage of high-frequency, high-resolution US machines in present practice [8]. Although several studies have evaluated the positive predictive value (PPV) of each US feature described in the fourth edition [1,6], at present there are no studies showing the PPV of US features from the fifth edition BI-RADS and how these changes affect performance in predicting malignancy among breast masses. In addition, most of the studies evaluating the diagnostic performance of the US BI-RADS lexicon are based on data review of static images of the breast mass, which may allow for the selection of representative images of the lesions, but cannot visualize the lesion as a whole.

Based on this, we evaluated the PPV of the grayscale US descriptors in the fifth edition of BI-RADS and investigated the effect of these changes by comparing the diagnostic performance of the fourth and fifth editions. For a more precise lesion characterization, we used video clips of the breast mass recorded during US examinations in data acquisition.

Materials and Methods

This prospective study has been approved by the Institutional Review Board of each of the five institutions participating in this study. Informed consent was obtained from patients for inclusion and the recording of video images during US examination. Signed, informed consent was obtained from all patients prior to US-guided biopsy, vacuum-assisted excision, or surgical procedures.

Patients

A multicenter database including five institutions of the Korean Society of Breast Imaging consists of women who had given consent for recording video clips of a specific breast mass during US examination. From the database, patients from October 2014 to November 2014 fulfilling the following criteria have been included in this study: (1) patients who had undergone surgery or US-guided vacuum assisted excision, (2) where percutaneous US-guided biopsy was performed, (3) with benign biopsy results, and patients had been followed for more than 2 years showing stability or decreased extent prior to the examination performed during the study period, and (4) with typically benign features described in the BI-RADS US lexicon in the "special cases" section [8], such as simple cyst, clustered microcysts, complicated cysts, mass in or on skin, intramammary lymph nodes, postsurgical fluid collection and fat necrosis. Images of 83 breast lesions in 77 women fulfilling the criteria above were obtained. Among them, three breast lesions in three patients were excluded from this study due to the poor image quality, which may affect image analysis, and mass size exceeding the measurement range or the visible depth of the transducer. Finally, sonograms of 80 breast masses in 74 women were included in this study. The mean size of the breast masses for which sonograms were obtained was 19.1±10.3 mm (range, 5 to 42 mm). The mean age of the 74 women was 47.5±10.7 years (range, 20 to 70 years).

US and Image Acquisition

For image acquisition, various US machines equipped with high-frequency linear array transducers were used (iU22, Philips Medical Systems, Bothell, WA, USA; GE LOGIQ E9, GE Medical Systems, Milwaukee, WI, USA; Supersonic Imagine, Aix en Provence, France; EUB-8500, Hitachi Medical, Tokyo, Japan). Seven radiologists dedicated to breast imaging with at least 3 years of experience (range, 3 to 13 years) were involved in patient collection and image acquisition. When a breast mass was detected during US examinations, routine scanning protocols were followed included scanning of transverse and longitudinal images of the mass, with and without calipers used for size measurement. After obtaining static images, video clips were recorded by the US machine. Video clips started at the area of normal breast parenchyma surrounding the mass, including the entire mass during one-direction movement of the probe, and ended at the other end of the mass including normal breast parenchyma. Video clips were obtained in transverse and longitudinal planes, as with the static images. The radiologist

Table 1. Ultrasonographic descriptors used in image analysis

Fourth edition	Fifth edition
Shape	Shape
Oval	Oval
Round	Round
Irregular	Irregular
Orientation	Orientation
Parallel	Parallel
Non-parallel	Non-parallel
Margin	Margin
Circumscribed	Circumscribed
Indistinct	Indistinct
Angular	Angular
Microlobulated	Microlobulated
Spiculated	Spiculated
Lesion boundary	–
Abrupt interface	–
Echogenic halo	–
Echo pattern	Echo pattern
Anechoic	Anechoic
Hyperechoic	Hyperechoic
Complex	Complex cystic and solid
Hypoechoic	Hypoechoic
Isoechoic	Isoechoic
–	Heterogeneous[a]
Posterior acoustic features	Posterior features
Absent	Absent
Enhancement	Enhancement
Shadowing	Shadowing
Combined	Combined
Calcifications	Calcifications
No calcifications	No calcifications
Macrocalcifications	–
Calcifications in mass	Calcifications in mass
Calcifications out of mass	Calcifications out of mass
Calcifications in/out of mass	–
–	Intraductal calcifications[a]
Surrounding tissue	Associated features
Ductal changes	Ductal changes
Cooper's ligament changes	–
Edema	Edema
Architectural distortion	Architectural distortion
Skin thickening	Skin thickening
Skin retraction	Skin retraction

'-', not defined.
[a]newly defined in 5th edition.

who had performed real-time breast US examinations and recording of video clips of the breast mass selected the representative static images of the mass, which were stored in an exclusive storage device along with the video images. Four images of the breast masses (static transverse, static longitudinal, video transverse, and video longitudinal) were displayed, in order, using Microsoft PowerPoint 2010 for image analysis.

Image Analysis According to the US BI-RADS Lexicon

Two radiologists dedicated to breast imaging (Y.M.K. and J.H.Y.) with 7 and 14 years of experience retrospectively reviewed the static and video images of the breast masses. Review of the static and video images was done independently. During image review, the radiologists were blinded to the histopathology results, clinical information such as the presence of symptoms, mammography findings, and the image analysis results of the other radiologist. Images for a breast mass were analyzed on the same day, first using the descriptors of the fourth edition of BI-RADS [5], and next, using the descriptors of the fifth edition [8] as in Table 1. Vascularity from Color Doppler and elastography were not included in the analysis of this study, since this was a multicenter study for which various US machines with different software were used in image acquisition.

Table 2. Histopathologic diagnosis of the 80 breast masses

	Pathology result	No. of cases (%)
Benign (n=51)	Abscess	1 (2.0)
	Adenomyoepithelioma	1 (2.0)
	Diabetic mastopathy	2 (3.9)
	Duct ectasia	6 (11.8)
	Fat necrosis	2 (3.9)
	Fibroadenoma	15 (29.3)
	Fibroadenomatous hyperplasia	4 (7.8)
	Fibrocystic change	4 (7.8)
	Galactocele	2 (3.9)
	Hamartoma	3 (5.9)
	Intraductal papilloma	5 (9.8)
	Intramammary lymph node	1 (2.0)
	Lipoma	1 (2.0)
	Phyllodes tumor, benign	1 (2.0)
	Postoperative fibrosis	3 (5.9)
Malignant (n=29)	Ductal carcinoma, *in situ*	4 (13.9)
	Invasive ductal carcinoma	20 (69.0)
	Invasive lobular carcinoma	1 (3.4)
	Invasive cribriform carcinoma	1 (3.4)
	Mucinous carcinoma	2 (6.9)
	Tubular carcinoma	1 (3.4)

Table 3. Ultrasonographic descriptors from fourth and fifth edition of BI-RADS according to pathology results

		Fourth edition				Fifth edition			
		Benign (n=51)	Malignant (n=29)	PPV (%)	P-value	Benign (n=51)	Malignant (n=29)	PPV (%)	P-value
Shape	Oval	34 (66.7)	6 (20.7)	15.0	<0.001	33 (64.7)	4 (13.8)	10.8	<0.001
	Round	6 (11.8)	0	0.0		7 (13.7)	1 (3.4)	12.5	
	Irregular	11 (21.5)	23 (79.3)	67.6		11 (21.6)	24 (82.8)	68.6	
Orientation	Parallel	37 (72.5)	14 (48.3)	27.5	0.030	35 (68.6)	11 (37.9)	23.9	0.008
	Non-parallel	14 (27.5)	15 (51.7)	51.7		16 (31.4)	18 (62.1)	52.9	
Margin	Circumscribed	25 (49.0)	2 (6.9)	7.4	0.001	23 (45.1)	0	0.0	<0.001
	Indistinct	8 (15.7)	4 (13.8)	33.3		7 (13.7)	6 (20.7)	46.2	
	Angular	6 (11.8)	9 (31.0)	60.0		8 (15.7)	7 (24.1)	46.7	
	Microlobulated	7 (13.7)	4 (13.8)	36.4		8 (15.7)	4 (13.8)	33.3	
	Spiculated	5 (9.8)	10 (34.5)	66.7		5 (9.8)	12 (41.4)	70.6	
Lesion boundary	Abrupt interface	47 (92.2)	21 (72.4)	30.9	0.017	–	–	–	–
	Echogenic halo	4 (7.8)	8 (27.6)	66.7		–	–	–	
Echo pattern	Anechoic	2 (5.9)	0	0.0	0.263	2 (3.9)	0	0.0	0.098
	Hyperechoic	3 (5.9)	0	0.0		2 (3.9)	0	0.0	
	Complex	13 (25.5)	10 (34.5)	38.5		8 (15.7)	0	0.0	
	Hypoechoic	28 (54.9)	19 (65.5)	39.1		23 (45.1)	17 (58.6)	42.5	
	Isoechoic	5 (9.8)	0	0.0		3 (5.9)	0	0.0	
	Heterogeneous	–	–	–		13 (25.5)	12 (41.4)	48.0	
Posterior acoustic feature	Absent	29 (56.9)	10 (34.5)	25.6	0.002	22 (43.1)	11 (37.9)	33.3	<0.001
	Enhancement	10 (19.6)	4 (13.8)	28.6		14 (27.5)	1 (3.5)	6.7	
	Shadowing	8 (15.7)	2 (6.9)	20.0		11 (21.6)	4 (13.8)	26.7	
	Combined	4 (7.8)	13 (44.8)	76.5		4 (7.8)	13 (44.8)	76.5	
Calcification	No calcifications	47 (92.2)	13 (44.8)	21.7	<0.001	48 (94.1)	12 (41.4)	20.2	<0.001
	Calcifications in mass	2 (3.9)	11 (37.9)	84.6		3 (5.9)	13 (44.9)	81.3	
	Calcifications out of mass	0	0	0.0		0	1 (3.4)	100.0	
	Calcifications in/out of mass	0	5 (17.3)	100.0		–	–	–	
	Macrocal	2 (3.9)	0	0.0		–	–	–	
	Intraductal calcifications	–	–	–		0	3 (10.3)	100.0	
Surrounding tissue	No	41 (80.4)	12 (41.4)	22.6	0.004	38 (74.5)	10 (34.5)	20.8	0.003
	Ductal changes	5 (9.8)	6 (20.7)	54.5		5 (9.8)	6 (20.7)	54.5	
	Cooper's ligament changes	0	0	0.0		–	–	–	
	Edema	1 (2.0)	3 (10.3)	75.0		3 (5.9)	2 (6.9)	40.0	
	Architectural distortion	3 (5.8)	8 (27.6)	72.7		5 (9.8)	11 (37.9)	68.8	
	Skin thickening	1 (2.0)	0	0.0		0	0	0.0	
	Skin retraction	0	0	0.0		0	0	0.0	
Category	2	11 (21.6)	0	0.0	<0.001	10 (19.6)	0	0.0	<0.001
	3	12 (23.5)	1 (3.4)	7.7		11 (21.6)	1 (3.4)	8.3	
	4a	17 (33.3)	2 (6.9)	10.5		18 (35.3)	2 (6.9)	10.0	
	4b	7 (13.7)	6 (20.7)	46.2		9 (17.6)	6 (20.7)	40.0	
	4c	4 (7.8)	14 (48.3)	77.8		3 (5.9)	14 (48.3)	82.4	
	5	0	6 (20.7)	100.0		0	6 (20.7)	100.0	

Values are presented as number (%) unless otherwise indicated.
PPV, positive predictive value.

Each radiologist chose and recorded the most appropriate term in each descriptor for each breast mass. Final assessments were made for each mass using one of the assessment categories of BI-RADS [8]: category 2, benign; 3, probably benign; 4a, low suspicion for malignancy; 4b, intermediate suspicion for malignancy; 4c, moderate concern for malignancy; and 5, highly suggesting malignancy. If the two radiologists had different opinions regarding the terms of descriptors or final assessment, the final decision was made based on consensus of the radiologists.

Data and Statistical Analysis

Histopathology from US-guided core needle biopsy, US-guided vacuum-assisted biopsy, and surgery was considered the standard reference. Breast masses showing typically benign US features, or which had been confirmed with biopsy as benign showing stability or decreased size during follow-up of more than 2 years prior to the study period, were considered benign.

An independent two-sample t test was used for comparison of continuous variables. A chi-square or Fisher exact test was used for comparison of the categorical variables between benign and malignant masses. Diagnostic performance including sensitivity, specificity, PPV, negative predictive value (NPV), and accuracy were calculated, and compared using generalized estimating equation analysis. For statistical analysis, category 2 and 3 lesions were considered benign, while category 4a to 5 lesions were considered malignant. Area under the receiver operating characteristics curve (AUC) was calculated for each edition of BI-RADS, and compared. A chi-square test was used in comparing the proportion of final assessment categories, and McNemar's test in evaluating the consistency of final assessment categories between the two editions. The concordance rate was calculated with upper and lower movement rates to see the changes among final assessment categories between the fourth and fifth editions of BI-RADS. Statistical analysis was performed using SAS ver. 9.2 (SAS Institute Inc., Cary, NC, USA). All tests were two-sided, and P-values of less

than 0.05 were considered to have statistical significance.

Results

Of the 80 breast masses, 51 (63.8%) were benign and 29 (36.2%) were malignant. Of the 77 women included in this study, 41 (53.2%) had palpable breast masses, three (3.9%) had breast pain, three (3.9%) had bloody nipple discharge, and the remaining 30 (39.0%) were asymptomatic. Histopathologic diagnosis of the 80 breast masses are summarized in Table 2.

US descriptors in the fourth and fifth editions of BI-RADS were compared between benign and malignant masses (Table 3). In the fourth edition, irregular shape (79.3%), non-parallel orientation (51.7%), angular (31.0%), and spiculated (34.5%) margins, echogenic halo (27.6%), and combined posterior acoustic features (44.8%) showed significantly higher PPV (all P<0.05). In the fifth edition, irregular shape (82.8%), non-parallel orientation (62.1%), indistinct (20.7%), angular (24.1%), and spiculated (41.4%) margins, and combined posterior features (44.8%) showed a significantly higher PPV (all P<0.05). Descriptors within an echo pattern did not show significant differences between benign and malignant cases in either edition. The PPV of the final assessment categories in the fourth and fifth edition are as follows: (1) fourth edition: 0.0% for category 2, 7.7% for category 3, 10.5% for category 4a, 46.2% for category 4b, 77.8% for category 4c, and 100.0% for category 5, (2) fifth edition: 0.0% for category 2, 8.3%

Table 4. Comparison of final assessment categories between fourth and fifth edition of BI-RADS

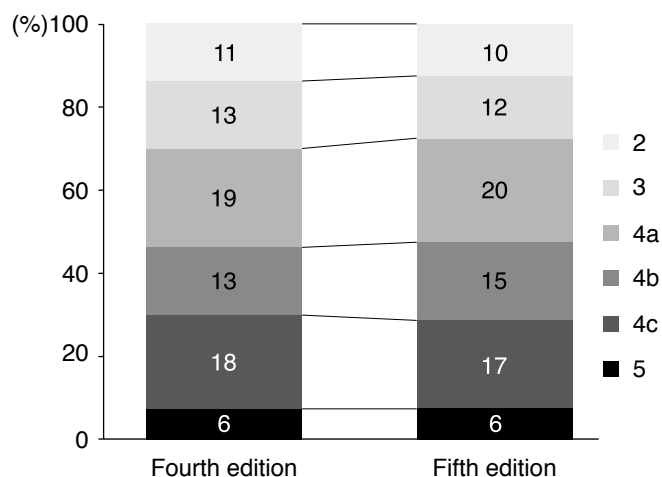| Category	Fourth edition	Fifth edition	P-value
2	11 (13.8)	10 (12.5)	0.815
3	13 (16.2)	12 (15.0)	0.828
4a	19 (23.8)	20 (25.0)	0.854
4b	13 (16.2)	15 (18.8)	0.677
4c	18 (22.5)	17 (21.2)	0.699
5	6 (7.5)	6 (7.5)	0.772

Values are presented as number (%).

Fig. 1. Box graph showing the changes among final assessment categories between the fourth and fifth edition of BI-RADS. McNemar test showed that there are no significant differences in the proportion of final assessment categories (P>0.999), suggesting consistency in final assessment between the fourth and fifth editions of BI-RADS.

Table 5. Comparison of diagnostic performances between the ultrasonographic final assessments in the fourth and fifth edition of BI-RADS

	Fourth edition	Fifth edition	P-value
Sensitivity	28/29 (96.6)	28/29 (96.6)	>0.999
Specificity	23/51 (45.1)	21/51 (41.2)	0.411
PPV	28/56 (50.0)	28/58 (48.3)	0.414
NPV	23/24 (95.8)	21/22 (95.5)	0.520
Accuracy	51/80 (63.8)	49/80 (61.3)	0.412
AUC (95% CI)	0.708 (0.632−0.785)	0.690 (0.613−0.765)	0.416

Values are presented as number (%) unless otherwise indicated.
PPV, positive predictive value; NPV, negative predictive value; AUC, area under the receiver operating characteristics curve; CI, confidence interval.

for category 3, 10.0% for category 4a, 40.0% for category 4b, 82.4% for category 4c, and 100.0% for category 5, respectively. In a comparison of the proportion of final assessment categories between the fourth and fifth editions of BI-RADS, no significant differences were seen between each edition of BI-RADS for each final assessment category (all P>0.05) (Table 4). McNemar's test showed that there are no significant differences in the proportion of final assessment categories (P>0.999) (Fig. 1). The concordance rate was 76.3% (61 of 80), with a 10% (8 of 80) upper movement rate and 13.8% (11 of 80) lower movement rate.

Diagnostic performance of the final assessment category were as follows (Table 5): for the fourth edition of BI-RADS, sensitivity 96.6%, specificity 45.1%, PPV 50.0%, NPV 95.8%, accuracy 63.8%; for the fifth edition of BI-RADS, sensitivity 96.6%, specificity 41.2%, PPV 48.3%, NPV 95.5%, and accuracy 61.3%, respectively. No significant differences were seen between the diagnostic indices between the two editions. The AUC was higher in the fourth edition, 0.708 to 0.690, but without significant differences (P=0.416).

Discussion

The fifth edition of BI-RADS is an extension of the fourth edition [8]; however, several changes are noticeable. First, the lesion boundary descriptor has been eliminated. As described in a prior study [2], the lesion boundary is not a major feature category, unlike shape or margin. This descriptor has been excluded from the US lexicon since an echogenic transition zone can be present in both malignant masses as well as benign abscesses [8]. In our study, malignant masses had significantly higher rates of echogenic halo than abrupt interfaces, and a higher PPV was also seen with echogenic halo (Table 3). On the other hand, approximately 72.4% of malignant masses had abrupt interfaces, which supports the description in the fifth edition of BI-RADS that the absence of echogenic halo is common

and considered to be of no diagnostic significance [8]. Second, "heterogeneous" echo pattern has been added in the fifth edition of BI-RADS, defined as a mixture of echogenic patterns within a solid mass, and although it has been reported to have little prognostic value in differentiating between benign and malignant masses, this feature may elevate the suspicion for malignancy, especially when seen with non-circumscribed margins and irregular shape [8]. In our study, malignant lesions showed significantly higher rates of a heterogeneous echo pattern compared to benign lesions, and the PPV of this feature was 48.0%. Interestingly, when compared to the analysis according to the fourth edition of BI-RADS, the PPV of complex cystic and solid masses in the fifth edition was 0.0%, remarkably decreased from the 38.5% in the fourth edition. Of the 29 malignant lesions, 34.5% were classified as complex echo in the fourth edition while 41.4% were classified as heterogeneous in the fifth edition and 0.0% as complex cystic and solid echo pattern (Table 3). Two reasons may provide an explanation for these results. First, although the two radiologists involved in image analysis were well aware of the definitions of the BI-RADS US lexicon, using them in daily practice, they are not accustomed to using "heterogeneous" echo pattern, which they may have mixed up with complex echo patterns. Second, this study is based on a retrospective review of still or video images from five institutions, among which US machines vary, which may have affected the results. Further prospective studies using real-time sonograms may be able to validate the true significance of a "heterogeneous" echo pattern among breast lesions.

Several studies have proven the fourth edition of the BI-RADS lexicon efficient and useful in describing and classifying breast lesions detected on breast US [1,2,6,9,10]. Although rather minor changes have been made in the fifth edition, this may affect the diagnostic performance of BI-RADS in some way, but at present, there are no studies evaluating the differences between the fourth and fifth editions of BI-RADS since the latter has been released. Similar values were observed when comparing diagnostic indices such as sensitivity, specificity, PPV, NPV, and accuracy (Table 5). AUC was 0.708 (95% confidence interval [CI], 0.632 to 0.785) for the fourth edition and 0.690 (95% CI, 0.613 to 0.765) for the fifth edition, without significant differences. Also, in a comparison of the proportion of final assessment categories between the fourth and fifth editions of BI-RADS, no significant differences were seen between each final assessment category or the proportion of final assessment categories, suggesting consistency in final assessment between the fourth and fifth editions of BI-RADS. During image analysis, we have come across cases that had different analysis results among descriptors that did not change between the two editions, such as shape, orientation, or margin.
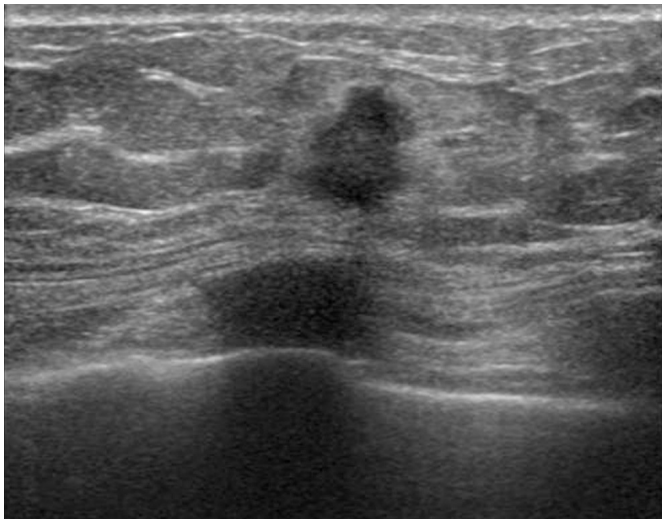
A
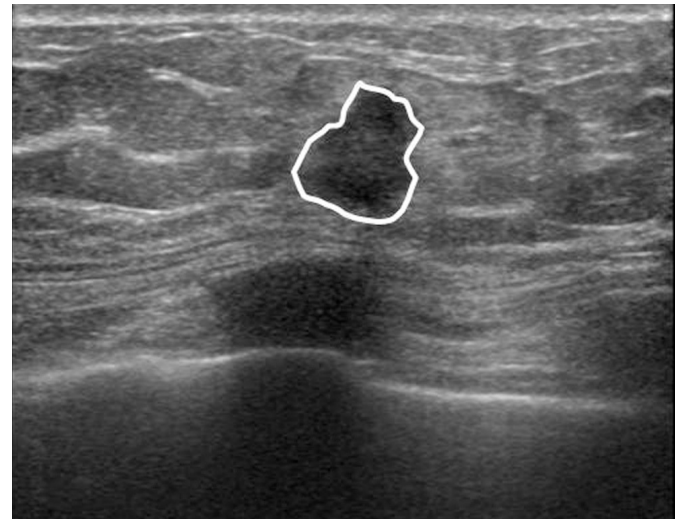
B
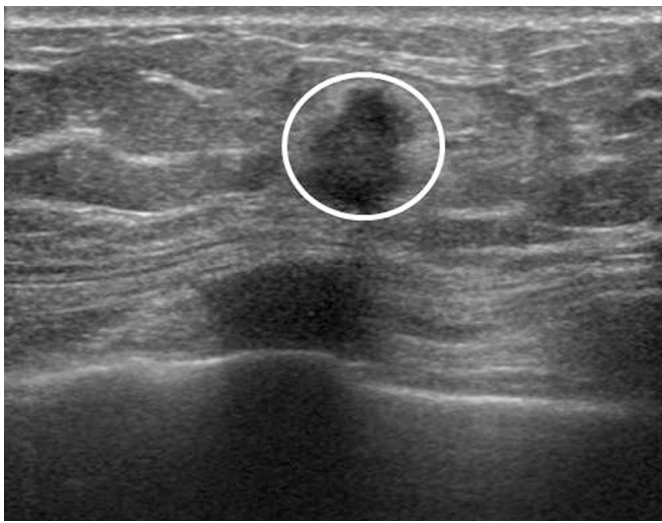
C

Fig. 2. Differences in image analysis between the fourth and fifth edition of BI-RADS influenced by elimination of lesion boundary descriptor on ultrasonography.
For a 8-mm breast mass was seen on ultrasonography (A), image analysis of this mass was irregular shape, not parallel orientation, angular margin, hypoechoic echo pattern and echogenic halo based on the fourth edition (B), but oval shape, parallel orientation, microlobulated margin, and heterogeneous echo pattern (C) when applying the fifth edition.

Exclusion of the lesion boundary descriptor seemed to be the cause of the discrepancy between the two editions, as described in Fig. 2. Although changes between the two editions may affect image analysis for individual descriptors, based on our results, we conclude that they have minimal effect in deciding upon a final assessment category for a breast lesion, and the fifth edition of the BI-RADS lexicon is as feasible and useful as the fourth edition in differentiating various breast masses detected on US.

One of the strong points of this study is that we used both the representative static images and the video clips including the entire breast mass recorded in real-time examinations during image analysis. Previously reported studies regarding the performance of the BI-RADS lexicon have been based on analysis of selected static images [1,2,6,9,10], and even if the images selected were those which were considered most representative, information is rather limited when compared to real-time visualization. Also, when considering the fair to moderate observer agreement regarding US features [7,11], opinion among observers may differ during image selection. To minimize the effect of observer variability that can affect our results and to approximate the real-time US examination environments of daily practice, we used video clips starting at the area of normal breast parenchyma surrounding the mass, including the entire mass during one-direction movement of the probe, and ending at the other end of the mass visualizing normal breast parenchyma. In contrast, the PPV of category 3 in the fourth and fifth edition, respectively, of our study was 7.7% and 8.3%, and that of category 4a was 10.5% and 10.0%. These rates are higher than the PPVs recommended for categories 3 and 4a by BI-RADS [8].

The limited number of cases included in this study may have been the cause of the higher PPV; PPV was calculated among less than 20 breast masses with category 3 or 4a assessment. Furthermore, this study was based on image analysis of pathologically confirmed breast masses that were enrolled after biopsy or excision, and since breast masses with clinical or radiological suspicion for malignancy are more prone to undergo biopsy, this may have been the cause for higher PPVs.

There are several limitations to this study. First, of the 51 benign masses, only lesions that had been diagnosed using biopsy were included. The false-negative rate of 14-gauge core needle biopsy has been reported to be approximately 2.5% in the literature [12,13], and the results of our study might have been different if all lesions included had been surgically confirmed. Second, elastography was not applied in lesion assessment. The fifth edition of the BI-RADS US lexicon includes elastography features in lesion description [8], as many institutions are equipped with US elastography devices. However, it is mentioned that the addition of elastography features in the US lexicon does not support its use in the characterization of breast lesions detected on US; rather, it is more likely for the purposes of acquiring a database to validate the true usefulness of elastography in lesion characterization, and therefore, it must not overrule the assessment of grayscale US features [8]. Third, the radiologists who reviewed the video clips were blinded to the mammography results. Interpretation of breast US and mammography during everyday practice is not done separately, but lesion correlation and integration of the features seen on the two imaging modalities are essential, which leads to a single final assessment regarding the abnormality detected. However, the main purpose of our study was to evaluate the PPV and performance of the fifth edition of the BI-RADS US lexicon; therefore, having knowledge of the mammographic features may have interfered with classifying US imaging features during feature analysis. Lastly, lesions showing stability over the follow-up period or typically benign features as described in the 'special cases' section of BI-RADS were included in this study based on imaging features only, without further pathologic confirmation. As different radiologists may not agree with the decision that these lesions can be considered benign based on imaging features only, they were included in this study based on the consensus of the seven radiologists involved in patient enrollment and image acquisition of this study, all agreeing that further biopsy of those lesions is not necessary when considering the typically benign features.

In conclusion, based on the results of our study, the fifth edition of the BI-RADS US lexicon shows comparable performance to the fourth edition, and can be useful in the differential diagnosis of breast masses using US.

ORCID: Jung Hyun Yoon: http://orcid.org/0000-0002-2100-3513; Min Jung Kim: http://orcid.org/0000-0003-4949-1237; Hye Sun Lee: http://orcid.org/0000-0001-6328-6948; Sung Hun Kim: http://orcid.org/0000-0003-4478-9720; Ji Hyun Youk: http://orcid.org/0000-0002-7787-780X; Sun Hye Jeong: http://orcid.org/0000-0002-3309-0693; You Me Kim: http://orcid.org/0000-0001-5807-9012

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

1. Costantini M, Belli P, Lombardi R, Franceschini G, Mule A, Bonomo L. Characterization of solid breast masses: use of the sonographic breast imaging reporting and data system lexicon. J Ultrasound Med 2006;25:649-659.
2. Kim EK, Ko KH, Oh KK, Kwak JY, You JK, Kim MJ, et al. Clinical application of the BI-RADS final assessment to breast sonography in conjunction with mammography. AJR Am J Roentgenol 2008;190:1209-1215.
3. Rahbar G, Sie AC, Hansen GC, Prince JS, Melany ML, Reynolds HE, et al. Benign versus malignant solid breast masses: US differentiation. Radiology 1999;213:889-894.
4. Shin HJ, Ko ES, Yi A. Breast cancer screening in Korean woman with dense breast tissue. J Korean Soc Radiol 2015;73:279-286.
5. American College of Radiology. Breast Imaging Reporting and Data System: BI-RADS Atlas. 4th ed. Reston, VA: American College of Radiology, 2003.
6. Hong AS, Rosen EL, Soo MS, Baker JA. BI-RADS for sonography: positive and negative predictive values of sonographic features. AJR Am J Roentgenol 2005;184:1260-1265.
7. Lee HJ, Kim EK, Kim MJ, Youk JH, Lee JY, Kang DR, et al. Observer variability of Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound. Eur J Radiol 2008;65:293-298.
8. D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. ACR BI-RADS Atlas, Breast Imaging Reporting and Data System. 5th ed. Reston, VA: American College of Radiology, 2013.
9. Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. Radiology 2006;239:385-391.
10. Sohn YM, Kim MJ, Kwak JY, Moon HJ, Kim SJ, Kim EK. Breast ultrasonography in young Asian women: analyses of BI-RADS final assessment category according to symptoms. Acta Radiol 2011;52:35-40.
11. Park CS, Lee JH, Yim HW, Kang BJ, Kim HS, Jung JI, et al. Observer

agreement using the ACR Breast Imaging Reporting and Data System (BI-RADS)-ultrasound, first edition (2003). Korean J Radiol 2007;8:397-402.

12. Youk JH, Kim EK, Kim MJ, Oh KK. Sonographically guided 14-gauge core needle biopsy of breast masses: a review of 2,420 cases with long-term follow-up. AJR Am J Roentgenol 2008;190:202-207.

13. Youk JH, Kim EK, Kim MJ, Kwak JY, Son EJ. Analysis of false-negative results after US-guided 14-gauge core needle breast biopsy. Eur Radiol 2010;20:782-789.