

# BI-RADS Category 4a Lesions of the Breast: Which US Features are More Predictive of Malignancy?

Jung Hyun Yoon, MD<sup>1,2</sup>, Eun-Kyung Kim, MD, PhD<sup>1</sup>, Min Jung Kim, MD, PhD<sup>1</sup>, Hee Jung Moon, MD, PhD<sup>1</sup>

<sup>1</sup>Department of Radiology, Research Institute of Radiological Science, Yonsei University, College of Medicine

<sup>2</sup>Department of Radiology, CHA Bundang Medical Center, CHA University, College of Medicine

**Background:** The purpose of our study was to evaluate which among the various ACR BI-RADS US descriptors predicted malignancy among breast lesions assessed as category 4a.

**Materials and Methods:** From January 2004 to September 2005, 494 breast lesions in 475 women which were assessed as category 4a on US, which had undergone biopsy/excision and follow-up were included in this study. Two radiologists retrospectively analyzed the US features of the 494 breast lesions based on the BI-RADS lexicon. Clinical features of the patients and US descriptors were analyzed according to histopathology results.

**Results:** Of the 494 category 4a lesions, 450 (91.1%) were benign and 44 (8.9%) were malignancy. PPV of category 4a lesions in this study was 8.9%. Among the US features, complex echoic features were more significantly seen in malignancy than benign, 15.9% to 7.5% ( $p = 0.043$ ). None of the US descriptors showed significant association to malignancy among the category 4a lesions on multivariate analysis.

**Conclusion:** Complex echoic features on US were more significantly seen in malignant lesions with category 4a assessment. Although none of the minor suspicious US features significantly predicted malignancy among category 4a lesions, usage of BI-RADS lexicon and assessment proves to be effective and consistent.

**Index words:** Breast; Ultrasound; BI-RADS

## Introduction

With the advance in image resolution and software, the usage of ultrasonography (US) in the breast has exponentially increased nowadays, mostly used in conjunction to mammography for both lesion detection and characterization (1-3). Although the widely spread screening using US can detect breast cancer at an early stage, this can lead to increased recommendations of

biopsies on benign breast lesions. As the usage of US in the breast has increased, the American College of Radiology (ACR) has developed a Breast Imaging Reporting and Data System (BI-RADS) for US (4), contributing in standardization of lesion description and classification. Many studies have proved this system as reliable, efficient and consistent in lesion description and final assessment, showing high predictive values (2, 5-9). This lexicon has various US descriptors, which in combination leads to a BI-RADS final assessment and recommendation. While the more benign appearing BI-RADS final assessment category 2 and 3 lesions are conservatively managed with follow-up, category 4 lesions are recommended for biopsy. Various clinical fac-

Corresponding to: Eun-Kyung Kim, M.D., Ph.D., Department of Radiology, Research Institute of Radiological Science, Yonsei University, College of Medicine, 50 Yonsei-ro, Shinchon-dong, Seodaemun-gu, Seoul 120-752, Korea.

Tel. 82-2-2228-7400 Fax. 82-2-393-3035 E-mail: ekkim@yuhs.ac

tors have been known to have effect on final assessments (9), and these clinical factors along with the combination of US descriptors can significantly increase or reduce unnecessary benign biopsies on solid breast lesions detected on screening US, in other words, many category 4a lesions may be assessed as category 3. Although a recent report evaluated the clinical factors which had effect on category 4 lesions (9), little has been investigated on how various US descriptors actually predict malignancy among category 4a lesions. Therefore, the purpose of our study was to evaluate which among the various US descriptors predicted malignancy among breast lesions assessed as category 4a.

## Materials & Methods

### *Patients*

This study was conducted retrospectively, and informed consent from patients was waived. Informed consent was obtained from each patient prior to biopsy procedures.

From January 2004 to September 2005, 1,150 breast lesions which were detected on US, assessed as category 4 or 5, had undergone pathologic evaluation at our institution. Among them, 647 (56.3%) lesions in 615 women were assessed as category 4a. Among them, 91 lesions were excluded because they had not undergone biopsy. Thirty lesions were excluded because although they were diagnosed as benign after biopsy, they had not been followed with imaging for at least 2 years afterwards. Also, 32 lesions which were diagnosed as high-risk lesions such as papillomas or phyllodes tumors which had not been pathologically confirmed with surgery were excluded. Finally, 494 breast lesions in 475 women which were assessed as category 4a on US were included in this study.

### *Imaging Methods, Biopsy Procedures, and Follow-up*

US examinations and subsequent US-guided biopsy procedures were performed by 1 of 7 dedicated breast radiologists with 1–11 years of experience. High-resolution breast US was performed using 5–12 MHz linear array transducers (HDI 3000 or 5000, Philips Advanced Technology Laboratories, Bothell, WA; Logic 9, GE Medical Systems, Milwaukee, WI, USA). Compound

imaging was performed in all examinations using the HDI 5000 machines. All US examinations included bilateral whole breast evaluation and investigation of any clinically- or mammographically-detected abnormalities.

US-guided biopsies were performed in all 494 lesions included in this study, prior to surgical or imaging-guided excision, if performed. US-guided core needle biopsies were performed with 14-gauge automated core biopsy needles with a 22-mm throw (SACN Biopsy Needle, Medical Device Technologies, Frankfurt, German) and automated guns (Promac 2.2L, Manan Medical Products, Northbrook, IL, USA), with free-hand technique. Vacuum-assisted biopsies were performed with 8-gauge or 11-gauge probes (Mammotome, Biopsys/Ethicon Endo-Surgery, Cincinnati, OH, USA). Local anesthesia was routinely applied before biopsy.

Patients with breast lesions diagnosed as benign on biopsy which showed imaging-pathologic concordance were recommended to follow-up US after 6 months, both mammography and US after 1 year, and yearly examinations thereafter (9).

### *BI-RADS Classification & Image Review*

All radiologists involved in US examinations were fully aware of the ACR BI-RADS of breast US, and BI-RADS lexicon and recommendations were used in radiologic reports. Two radiologists, each with 9 and 15 years of experience in breast imaging, retrospectively analyzed the US features of the 494 breast lesions based on the BI-RADS lexicon. Both radiologists were blinded to the clinical or pathologic information of the patient during image analysis.

Suspicious US features were defined based on our experience (2), ACR BI-RADS (4), and the criteria established by Stavros et al. (1). Suspicious US features included were irregular shape; spiculated, microlobulated, or angular margin; microcalcifications; non-parallel orientation; ductal extension; complex echogenicity; and posterior shadowing. Based on the study of Kim et al. (2), suspicious US features were further divided into major; irregular shape, spiculated margins and microcalcifications, and minor; round shape, microlobulated/in-distinct/angular margins, non-parallel orientation, duct extension, complex echogenicity, and posterior shadowing. Also, since thick echogenic halos are mostly

known to represent hypoechoic spiculations which from the lack of US resolution cannot be detected (1), lesion boundary was also excluded from analysis. Category 4a included in this study was defined as lesions showing 1 to 3 minor suspicious US features, as in a recent report (9).

**Table 1.** Pathology Results of the 494 Breast Lesions Assessed as Category 4a

|                              | Final Pathology (No.)                        |
|------------------------------|----------------------------------------------|
| Benign<br>(n = 450, 91.1%)   | Abscess, acute or chronic (12)               |
|                              | Adenosis (16)                                |
|                              | Atypical ductal hyperplasia (2)              |
|                              | Ductal epithelial hyperplasia (11)           |
|                              | Fibroadenoma (79)                            |
|                              | Fibrocystic change (131)                     |
|                              | Fibroadenomatous hyperplasia (53)            |
|                              | Galactocele (10)                             |
|                              | Intraductal papilloma (37)                   |
|                              | Phyllodes tumor, benign (17)                 |
|                              | Sclerosing adenosis (19)                     |
|                              | Stromal fibrosis (63)                        |
| Malignancy<br>(n = 44, 8.9%) | Apocrine carcinoma (1)                       |
|                              | Ductal carcinoma, in situ (8)                |
|                              | Invasive cribriform carcinoma (2)            |
|                              | Invasive ductal carcinoma (24)               |
|                              | Invasive lobular carcinoma (1)               |
|                              | Mucinous carcinoma (1)                       |
|                              | Papillary carcinoma (3)                      |
|                              | Phyllodes tumor, borderline or malignant (2) |
|                              | Others (2)                                   |

### Statistical Analysis

Pathologic diagnosis based on the US-guided core needle biopsy, vacuum-assisted percutaneous biopsy/excision, or surgical excision was set as standard reference. Positive predictive value (PPV) of category 4a lesions included in this study was calculated. Student's *t*-test was used for continuous variables.  $\chi^2$  test was used to compare the differences among each US descriptor among malignant lesions. Multivariate logistic regression analysis with odds ratio (OR) estimates and 95% confidence interval (CI) was also calculated to see the effects of each US descriptor in predicting malignancy among category 4a lesions. Statistical analyses were performed using Statistical Package for the Social Sciences (SPSS version 18.0, Chicago, IL).  $P < 0.05$  was considered statistically significant.

### Results

Of the 494 breast lesions assessed as category 4a included in this study, 450 (91.1%) were benign and 44 (8.9%) were malignancy. Table 1 summarizes the final pathology results of both benign and malignant lesions. PPV of category 4a was 8.9%. Mean age of the patients included was 44.2 years (range: 20–81 years). Mean size of the 517 breast lesions included was 12.6 mm (range, 5–20 mm). Mean age of the patients who were diagnosed as malignancy was significantly older than

**Table 2.** Comparison of US Descriptors with Malignancy in 494 Breast Lesions Assessed as Category 4a

| US Features        | Descriptors     | Pathology (%)    |                     |            | P     |
|--------------------|-----------------|------------------|---------------------|------------|-------|
|                    |                 | Benign (n = 450) | Malignancy (n = 44) | Total      |       |
| Shape              | Oval            | 392 (87.1)       | 34 (77.3)           | 426 (86.2) | 0.054 |
|                    | Round           | 45 (10.0)        | 7 (15.9)            | 52 (10.5)  |       |
|                    | Lobular         | 13 (2.9)         | 3 (6.8)             | 16 (3.3)   |       |
| Margin             | Circumscribed   | 050 (11.2)       | 5 (11.4)            | 55 (11.1)  | 0.293 |
|                    | Microlobulated  | 280 (62.2)       | 24 (54.5)           | 304 (61.5) |       |
|                    | Angular         | 33 (7.3)         | 2 (4.6)             | 35 (7.2)   |       |
|                    | Indistinct      | 087 (19.3)       | 13 (29.5)           | 100 (20.2) |       |
| Echo pattern       | Hyper/isoechoic | 161 (35.8)       | 11 (25.0)           | 172 (34.8) | 0.043 |
|                    | Hypoechoic      | 255 (56.7)       | 26 (59.1)           | 281 (56.9) |       |
|                    | Complex         | 34 (7.5)         | 7 (15.9)            | 41 (8.3)   |       |
| Orientation        | Parallel        | 359 (79.8)       | 33 (75.0)           | 392 (79.4) | 0.581 |
|                    | Non-parallel    | 091 (20.2)       | 11 (25.0)           | 102 (20.6) |       |
| Posterior Features | No              | 406 (90.2)       | 38 (86.4)           | 444 (89.9) | 0.499 |
|                    | Enhancement     | 42 (9.3)         | 6 (13.6)            | 48 (9.7)   |       |
|                    | Shadowing       | 02 (0.4)         | 0 (0.0)             | 2 (0.4)    |       |

those who had benign lesions,  $46.7 \pm 9.3$  years to  $43.7 \pm 8.6$  years ( $p = 0.029$ ). Mean size of the breast lesions which were diagnosed as malignancy did not show significant differences when compared to benign ones,  $14.7 \pm 9.6$  mm to  $11.7 \pm 9.6$  mm ( $p = 0.054$ ). Of the 494 patients, 114 (23.1%) had symptoms such as palpability, pain or nipple discharge, and 380 (76.9%) were asymptomatic. Symptomatic lesions were more significantly observed in malignancy ( $n=94$ , 20.9%) than benign ( $n=20$ , 45.5%) ( $p < 0.001$ ).

Comparison of US descriptors among the 494 breast lesions assessed as category 4a is summarized in Table 2, 3. Among the US features, complex echoic features were more significantly seen in malignancy than benign, 15.9% to 7.5% ( $p = 0.043$ ). Also, none of the US descriptors showed significant association to malignancy on multivariate analysis.

## Discussion

US shows excellent performances in the differentiation between benign and malignant breast lesions (1, 10, 11), plus the benefits of no radiation hazard, less patient discomfort during examinations, and simplicity of subsequent US-guided biopsy procedures if needed. Also, as proposed by several studies, US can be effectively used in screening of women with dense breasts (12–14). In particular, as the incidence of breast cancer has continuously increased in Korea, women who actively participate in breast cancer screening have also increased. Since most women of the Korean population have dense breast on mammography, and more young women are participating in screening of the breast, US examinations of the breast is now a widely spread, common study (15), performed by clinicians or radiologists of various experience.

The US ACR BI-RADS descriptors have been reported to have high diagnostic accuracy, positive and negative predictive values in differentiating between benign and malignant lesions (2, 6, 8), but still nodules of equivocal features exist. Although recommendations for category 4a is biopsy, the expected malignancy of category 4a lesions are 3–10% (4), which means that over 90% of the patients with category 4a lesions undergo invasive biopsy procedures for benign diagnosis. Efforts have been made to evaluate the precise features that may help in reducing these benign biopsies, mostly using additional techniques to conventional US such as elastography, but much more is still left to investigate. In this study, we analyzed the US descriptors of ACR BI-RADS among category 4a lesions, and tried to evaluate which of the minor suspicious features in category 4a lesions were more predictive of malignancy among US examinations only.

Complex echogenic features are defined as the mixture of solid and cystic contents within a breast lesion, which is well-known to be associated with papillary lesions of the breast (16, 17). Pathologically, intraductal papillomas frequently have internal necrosis or hemorrhage (18), which contribute to the complex echogenic features. Imaging features are similar between benign intraductal papillomas and papillary carcinomas, since malignant papillary carcinomas tend to show relatively benign US features, such as round to oval shape and well-defined margins (19). Differential diagnosis based on imaging features only may be difficult in papillary lesions of the breast, and this may explain why complex echoic features showed significant differences between benign and malignant lesions among the category 4a lesions.

As in the results of our study, none of the minor suspicious US features were significantly associated to malignancy.

**Table 3.** Comparison of US Descriptors Using Multivariate Logistic Regression Analysis

| US Features        | $\beta$ | OR    | 95% CI |       | P     |
|--------------------|---------|-------|--------|-------|-------|
|                    |         |       | Lower  | Upper |       |
| Shape              | 0.520   | 1.683 | 0.940  | 3.013 | 0.080 |
| Margin             | -0.088  | 0.916 | 0.696  | 1.205 | 0.531 |
| Echo Pattern       | 0.307   | 1.360 | 0.989  | 1.869 | 0.580 |
| Orientation        | 0.002   | 1.002 | 0.466  | 2.157 | 0.995 |
| Posterior Features | 0.275   | 1.317 | 0.538  | 3.219 | 0.546 |

nancy among category 4a lesions, but when combined together, the final assessments were within the recommendations of ACR BI-RADS; PPV of category 4a lesions in this study was 8.9%. This further supports the efficacy and reproducibility of ACR BI-RADS lexicon, assessments, and recommendations along with the many previous reports (2, 6-8). In addition, although we had defined category 4a lesions as those showing 1 to 3 minor suspicious US features (9), to the present date, there are no studies on a systematic approach to category 4 lesions of the breast. Minor suspicious US features in particular may be more predictive of malignancy in combinations, and further studies are anticipated with a large study population to figure out which sort of combination of minor suspicious US features may be fit in category 4a assessment.

There are some limitations to this study. First, 41.1% (185 of 450) of the benign lesions were diagnosed with US-guided core needle biopsy only, and not with excision. Although false-negative results are reported to be approximately 2.5% with US-guided core needle biopsy (20), the lesions included in our study were also those which were regularly followed with US, and did not show changes during follow-up. Second, observer variability may have existed between the 2 radiologists who retrospectively reviewed the images of the breast lesions.

In conclusion, complex echoic features on US were more significantly seen in malignant lesions with category 4a assessment. Although none of the minor suspicious US features significantly predicted malignancy among category 4a lesions, usage of BI-RADS lexicon and assessment proves to be effective and consistent.

## References

1. 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
2. Kim EK, Ko KH, Oh KK, 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. Benson SR, Blue J, Judd K, Harman JE. Ultrasound is now better than mammography for the detection of invasive breast cancer. *Am J Surg* 2004;188:381-385
4. ACR BI-RADS : breast imaging system : breast imaging atlas : mammography, breast ultrasound, magnetic resonance imaging. 4th ed. ed. Reston, Va.: American College of Radiology, 2003: 1 v. (loose-leaf)
5. 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; quiz 661
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. 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
8. Lee HJ, Kim EK, Kim MJ, et al. Observer variability of Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound. *Eur J Radiol* 2008;65:293-298
9. Yoon JH, Kim MJ, Moon HJ, Kwak JY, Kim EK. Subcategorization of ultrasonographic BI-RADS category 4: positive predictive value and clinical factors affecting it. *Ultrasound Med Biol* 2011;37:693-699
10. Rahbar G, Sie AC, Hansen GC, et al. Benign versus malignant solid breast masses: US differentiation. *Radiology* 1999;213:889-894
11. Jackson VP. Management of solid breast nodules: what is the role of sonography? *Radiology* 1995;196:14-15
12. Kolb TM, Lichy J, Newhouse JH. Occult cancer in women with dense breasts: detection with screening US--diagnostic yield and tumor characteristics. *Radiology* 1998;207:191-199
13. Buchberger W, DeKoekkoek-Doll P, Springer P, Obrist P, Dunser M. Incidental findings on sonography of the breast: clinical significance and diagnostic workup. *AJR Am J Roentgenol* 1999;173:921-927
14. Kaplan SS. Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue. *Radiology* 2001;221:641-649
15. 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
16. Lam WW, Chu WC, Tang AP, Tse G, Ma TK. Role of radiologic features in the management of papillary lesions of the breast. *AJR Am J Roentgenol* 2006;186:1322-1327
17. Brookes MJ, Bourke AG. Radiological appearances of papillary breast lesions. *Clin Radiol* 2008;63:1265-1273
18. Han BK, Choe YH, Ko YH, Yang JH, Nam SJ. Benign papillary lesions of the breast: sonographic-pathologic correlation. *J Ultrasound Med* 1999;18:217-223
19. Kim TH, Kang DK, Kim SY, Lee EJ, Jung YS, Yim H. Sonographic differentiation of benign and malignant papillary lesions of the breast. *J Ultrasound Med* 2008;27:75-82
20. 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

## BI-RADS Category 4a 유방 병변 중 어떠한 초음파 기술어가 악성 병변을 예측하는가?

윤정현<sup>1, 2</sup> · 김은경<sup>1</sup> · 김민정<sup>1</sup> · 문희정<sup>1</sup>

<sup>1</sup>연세의료원, 연세대학교 의과대학 영상의학과

<sup>2</sup>분당차병원, 차의과대학교 영상의학과

**목적:** 본 연구의 목적은 ACR BI-RADS의 초음파 기술어 중 어떠한 것이 category 4a의 병변 중 악성 병변을 예측할 수 있는지에 대해서 알아보고자 한다.

**대상 및 방법:** 본 연구는 2004년 1월부터 2005년 9월까지, 초음파로 category 4a로 진단받은 494개의 유방 병변을 가진 475명의 여성을 대상으로 하였다. 모든 유방 병변에서 조직검사, 국소절제술, 혹은 추적검사가 진행되었다. 두 명의 영상의학과 전문의가 후향적으로 494개의 유방 병변의 초음파 사진을 ACR BI-RADS의 기술어에 따라 분석하였다. 연구에 포함된 환자의 임상적 특징, 초음파 기술어와 최종 진단을 비교 분석하였다.

**결과:** 총 494의 유방 병변 중 450 (91.1%)개는 양성, 44 (8.9%)개는 악성으로 진단되었다. Category 4a의 양성예측도는 8.9%였다. 초음파 소견 중 혼합 에코성 병변에서 악성 병변이 양성 병변보다 더 유의하게 많이 나타났다 (15.9% 대 7.5%,  $p = 0.043$ ). 다변량 분석에서 어떠한 초음파 기술어도 category 4a의 병변 중 유의하게 악성도와 관련있는 것은 없었다.

**결론:** 혼합 에코성 병변은 양성 병변과 비교하여 악성 병변에서 유의하게 많이 나타났다. 초음파 기술어 중 category 4a 병변의 악성도를 유의하게 예측할 수 있는 것은 없었으나, BI-RADS 기술어를 사용한 병변의 최종 평가는 효율적이고 일관성 있는 것으로 나타났다.

**Index words:** Breast; Ultrasound; BI-RADS