A comparison between short-interval and regular-interval follow-up for BI-RADS category 3 lesions

Ruamsup S, Wiratkapun C, Wibulpolprasert B, Lertsithichai P

ABSTRACT

<u>Introduction</u>: The aim of this study was to compare the short-interval and regular-interval follow-up in women with Breast Imaging Reporting and Data System (BI-RADS) category 3 screen imaging studies. The image stability, rate of image-detected breast cancer and stage of cancer are studied.

<u>Methods</u>: Women who had BI-RADS 3 screen imaging studies (mammography and ultrasonography) conducted between the period January 2003 and December 2005 were retrospectively identified using the computerised database at the Department of Radiology, Ramathibodi Hospital, Thailand. Women who had known breast cancer status at two years after screening were included in the study and divided into two groups: short-interval (six months after screening) or regular-interval (one year after screening) follow-up. The two groups were compared in terms of the baseline clinico-radiologic characteristics and outcomes, including the image stability, image-detected breast cancer and the cancer stage at detection.

<u>Results</u>: A total of 10,086 women underwent screen imaging studies within the study period. Of these, 1,541 (15 percent) were categorised as BI-RADS 3. Only 1,036 women (67 percent) had follow-up images done six to 12 months after screening, and 846 (82 percent of 1,036 women) also had known cancer status two years after the screening. Breast cancer was noted in seven women (a positive predictive value of 0.7 percent). There were no significant differences between the two groups of women in terms of their baseline characteristics, image stability at the initial follow-up, the rate of image-detected breast cancer and the stage of cancer at detection.

<u>Conclusion</u>: There were no significant differences in the effectiveness of short-interval versus regular-interval follow-up in women with BI-RADS 3 screen imaging studies.

Keywords: BI-RADS category 3, breast cancer, regular-interval follow-up, screen imaging studies, short-interval follow-up

Singapore Med J 2010; 51(2): 120-125

Department of Radiology, Ramathibodi Hospital and Medical School, Mahidol University, 270 Rama 6 Road, Ratchatewi, Bangkok 10400, Thailand

Ruamsup S, MD
Fellow

Wiratkapun C, MD
Assistant Professor

Wibulpolprasert B, MD
Associate Professor

Department of Surgery

Lertsithichai P, MD
Associate Professor

Correspondence to:
Dr Panuwat Lertsithichai
Tel: (66) 2 201 1315
Fax: (66) 2 201 1316
Email: raplt@mahidol. ac.th

INTRODUCTION

Women with Breast Imaging and Data Reporting System category 3 (BI-RADS 3) imaging findings (mammography and ultrasonography) have an associated risk of a concurrent breast cancer of 2% or less.[1,2] The accepted recommendation in most guidelines for this group of women is short-interval imaging follow-up.[3,4] The "short-interval" is usually three to six months. The American College of Radiology recommends six-month imaging follow-up for at least two years.[4]

There have been many studies addressing the question of whether or not the short-interval follow-up recommendation for women with BI-RADS 3 imaging findings is appropriate. The rationale behind a short-interval follow-up is to detect existing malignancy at an early stage. Previous studies have never directly answered the question of whether short-term follow-up is more effective than regular-interval (annual) follow-up.[5] For example, many studies have addressed the question of the predictive value of BI-RADS 3 mammograms,[1,2,6] the addition of other imaging studies to the mammogram,[6] the comparison of the predictive values between various categories of mammograms or imaging studies,[1,8] or the accuracy of short-interval mammograms compared to first-performed screen or diagnostic mammograms,[9,10] which are all of indirect relevance to the question.

A direct approach to addressing the question would be to perform a comparative study of two groups of women with BI-RADS 3 on initial screen imaging, with one group undergoing short-interval follow-up and the other undergoing regular-interval follow-up, and to compare the two groups in terms of the imaging findings on follow-up, the positive predictive value for breast

cancer at two years post screening, image-detected breast cancer and the stage of breast cancer at detection. The objective of the present study was to perform such a comparative study.

METHODS

Imaging studies of women undergoing screening mammography and ultrasonography during the three-year period, between January 2003 and December 2005 at the Breast Diagnostic Centre, Ramathibodi Hospital, Thailand, were reviewed. The hospital's Research Ethics Committee approved the study. The choice between short- or regular-interval follow-up was most relevant for asymptomatic women, since imaging studies were the only available means of follow-up. Women who had breast imaging studies conducted for breast symptoms were therefore excluded. The breast imaging database of women with screen imaging studies reported as BI-RADS 3 were further searched for follow-up imaging studies that had been conducted during the subsequent 24 months. Biopsy and clinical examination results of any breast lesion found during the two-year follow-up period were obtained, as far as possible.

Breast cancer was determined to be present if the open or core needle biopsy or mastectomy specimens were positive for breast cancer. Breast cancer was determined to be absent if the histological examination revealed benign findings or a final imaging study at 24 months showed a BI-RADS 1, 2 or 3 lesion if biopsies were not performed. Women who did not complete the two-year follow-up were excluded from the study. Some women were also excluded if the breast cancer status at two years could not be determined. These included women with BI-RADS 4–5 at the 24th month of follow-up who did not undergo biopsy of the lesion. For each BI-RADS category, the positive predictive value (PPV) for breast cancer was defined as the ratio between the number of breast cancers and the total number of women with known cancer status for that particular BI-RADS category.

Mammography was performed using standard methods. Prior to the beginning of 2005, all mammograms were obtained using analogue machines (Lorad M-IV, Danbury, CT, USA; and Senographe DMR, GE, Milwaukee, WI, USA). Thereafter, almost all mammograms were obtained with a full-field digital mammography machine (Lorad Selenia, Hologic, Danbury, CT, USA). Simultaneously, sonograms (HDI 5000, Philips Ultrasound, Bothell, WA, USA) were obtained for almost all patients, with the exception of those with almost entirely fatty breasts. All the imaging studies were performed by experienced radiologists.

Screen imaging studies were defined as those performed for asymptomatic women or if the referring physicians marked the request forms as such. The screen, or index images were defined as the earliest screen images obtained for a particular woman within time period between January 2003 and December 2005. There may have been earlier images obtained prior to January 2003 for some patients, but these were excluded. Follow-up images were usually performed six months to one year after screening. The BI-RADS categories were not reported separately for mammograms and sonograms; the reported category was chosen as the highest of the two sets of images. The breast imaging information was complete for all the women.

Images were defined as stable if the initial follow-up images (the first set of short- or regular-interval follow-up images obtained after the index screening imaging) were categorised as BI-RADS 3, 2 or 1. The images were regarded as progressive if the initial follow-up images were categorised as BI-RADS 4 or 5. The initial follow-up image-detected breast cancer was defined as histology-proven breast cancer, subsequent to the categorisation of initial follow-up images as BI-RADS 4 or 5. All subsequently detected breast cancers were located in the same area as the lesions that were seen on the imaging studies.

Other data obtained from the medical records included age at the index imaging, breast density, size of the lesions seen on imaging, mammographic and sonographic findings, the results of a definitive pathological examination, pathological size and the tumour, node, metastasis (TNM) stage.

The primary comparison was between women with BI-RADS 3 screen images who underwent short-interval follow-up and those who underwent regular-interval follow-up imaging, and who were followed for at least two years after the initial screen. Due to the observational character of the present study, the choice of short- or regular-interval follow-up was left to the preferences of the patient and her referring physician. The outcomes to be compared were the rate of image stability, PPV for breast cancer at two years, the image-detected breast cancer rate at the initial follow-up, and the cancer stage.

The continuous variables were summarised using the mean ± SD or median (range), as appropriate. The categorical variables were summarised using counts and percentages. Statistical tests of differences in the continuous variables between the two independent groups were performed using a t-test or Wilcoxon rank-sum test, as appropriate. Differences in the categorical variables

Table I. Characteristics of all the patients screened.

Characteristic (n = 10,086, unless otherwise stated)	No. (%)
Age at initial screen imaging (years)	
Mean ± SD	51.1 ± 7.9
Median (Range)	51 (20–99)
Breast density (BI-RADS mammographic criteria)	
Dense (extreme and heterogeneous density)	8,640 (86)
Not dense (fibro-glandular and fatty)	1,446 (14)
BI-RADS category at screen imaging	
1	2,697 (26.7)
2	5,670 (56.2)
3	1,541 (15.3)
4A	26 (0.3)
4B	121 (1.2)
4C	11 (0.1)
5	20 (0.2)
Biopsy	
Conducted	251 (2.5)
Not conducted	9,835 (97.5)
Cancer status after two years follow-up*	
Cancer	78 (0.8)
No cancer	-
Low-risk lesion	6,768 (67.1)
High-risk lesion (LCIS, ALH, ADH)	16 (0.2)
Uncertain	3,224 (31.9)
Positive predictive value for cancer (n = 6,862)**	
BI-RADS 1	2/1,614 (0.1)
BI-RADS 2	19/3,974 (0.5)
BI-RADS 3	9/1,128 (0.8)
BI-RADS 4A	1/18 (5.6)
BI-RADS 4B	25/101 (24.8)
BI-RADS 4C	7/10 (70.0)
BI-RADS 5	15/17 (88.2)

* The cancer status was defined according to histological results obtained within two year after the screening. "No cancer" was also defined as BI-RADS 1–3 at two years after screening; otherwise the lesion was defined as "uncertain".
** Determined for women with known cancer status at two years.
SD: standard deviation; BI-RADS: Breast Imaging and Data Reporting System; LCIS: lobular carcinoma *in situ*; ALH: atypical lobular hyperplasia; ADH: atypical ductal hyperplasia

were tested using the chi-square test or Fisher's exact test, as appropriate. Statistical significance was defined as a two-sided test with a p-value of 0.05 or less. All statistical analyses were performed using Stata version 9 (Stata Corp, College Station, TX, USA).

RESULTS

A total of 10,086 women underwent a screen breast imaging study during the time period between January 2003 and December 2005. Of these, 1,541 (15%) women had BI-RADS 3 initial image results, and 6,862 (68%) had known cancer status at 24 months post screening. The presence of breast cancer was found in 78 (1.1%) of 6,862 women at 24 months post screening. Only 1,036 (67%) of 1,541 women with BI-RADS 3 screen images also had follow-up imaging studies performed at six to 12 months post screening. These women were the main subjects of the present study. Of these, 846 (82%) of 1,036 women had known cancer status at 24 months.

The clinico-radiologic characteristics of the women who underwent screening imaging studies are presented in Table I, along with the PPV for breast cancer of each BI-RADS category. The baseline and outcome comparisons between women with screen BI-RADS 3 images who underwent short-interval follow-up and those who underwent regular-interval follow-up are presented in Table II. As shown in Table II, there were no significant differences in the baseline characteristics and outcomes of the two groups. Only 30% of all the women with BI-RADS 3 index images underwent short-interval follow-up, while 70% underwent regular-interval follow-up. The PPV for breast cancer was 0.7% (six out of 846) at two years, and the initial follow-up images were able to detect 67% (four out of six) of these cancers. All the breast cancers detected were in the early-stage. Only 3% had interval progression, as seen on the initial follow-up images, i.e, there was 97% image stability.

For women with BI-RADS 3 screen images, the comparison between those with known and those with unknown cancer status at 24 months after the screening

Table II. Comparison of short-interval and regular-interval follow-up for women with BI-RADS 3 screen images.

Characteristic	No (%)			p-value[a]
	Short-interval follow-up (n = 368)	Regular-interval follow-up (n = 668)	All BI-RADS 3 (n = 1,036)	
Mean age ± SD (years)	49.5 ± 6.9	50.1 ± 7.2	49.9 ± 7.1	0.189
Dense breast	332 (90.2)	606 (90.7)	938 (90.5)	0.792
Biopsy	10 (2.7)	28 (4.2)	38 (3.7)	0.227
Low risk lesion	6 (1.6)	21 (3.1)	27 (2.6)	0.143
High risk lesion	1 (0.3)	2 (0.3)	3 (0.3)	0.937
Cancer	2 (0.5)	4 (0.6)	6 (0.6)	0.908
Interval progression (at initial short or regular follow-up)	10 (2.7)	23 (3.4)	33 (3.2)	0.524
PPV for breast cancer[b]	2/278 (0.7)	4/568 (0.7)	6/846 (0.7)	0.980
Rate of image detected cancer[b,c]	1/278 (0.4)	3/568 (0.5)	4/846 (0.5)	0.999
pTNM stage[b,d]				
0: DCIS	0/278 (0.0)	1/567 (0.2)	1/845 (0.1)	0.531
I	1/278 (0.4)	1/567 (0.2)	2/845 (0.2)	
IIA	0/278 (0.0)	1/567 (0.2)	1/845 (0.1)	
IIB	1/278 (0.4)	0/567 (0.0)	1/845 (0.1)	

[a] This is the p-value according to *t*-test, rank test, chi-square or Fisher's exact test, as appropriate.
[b] There were only 846 of 1,036 women with known cancer status at 24 months.
[c] Detected at initial short- or regular-interval follow-up.
[d] One patient was transferred to another hospital before operative treatment was conducted.
BI-RADS: Breast Imaging and Data Reporting System; SD: standard deviation; PPV: positive predictive value; pTNM: pathological tumour, node, metastasis; DCIS: ductal carcinoma *in situ*

is shown in Table III. The significant difference in the frequency of performed biopsies was expected, as a histological diagnosis was one of the criteria used for the determination of cancer status. The two biopsies in the unknown cancer status group were performed at other hospitals and the formal histological reports were not obtained. Significantly more women underwent short-interval follow-up in the unknown status group than in the known status group.

DISCUSSION

Many studies have examined the PPV of breast cancer for various BI-RADS categories determined within one to two years after the index imaging study.[1,2,6,8-10] All these studies have found the PPV for breast cancer of BI-RADS 3 lesions to be low, and that asymptomatic women with such images do not routinely require breast biopsies. Instead, short-interval imaging follow-up is often recommended.[11] However, the greater effectiveness of a short-term follow-up over regular-interval follow-up has never been proven.[2,5,9,11,12]

Previous studies have also not directly addressed the effectiveness of short-interval follow-up vs. regular-interval follow-up recommendations for women with BI-RADS 3 breast images. Since the aim of a shorter-interval follow-up is to detect existing breast cancers at an early stage and at an earlier point in time, the most valid study design would be to conduct a comparison of short- and regular-interval follow-up strategies in terms of the radiologic breast cancer detection rate and the stage of cancer at some appropriate time point (one or two years) after the index imaging studies.

The present study found no significant differences in the image stability (interval progression of the lesions or newly seen lesions), the PPV for breast cancer at two years, the rate of image-detected breast cancer at the initial follow-up and the stage of cancer at detection between women with BI-RADS 3 screen images who either underwent short-interval or regular-interval follow-up. Initial follow-up images were able to detect four out of six (67%) histology-proven breast cancers, a sensitivity rate that is similar to that found in previous studies.[9,10]

There are several possible explanations that are not necessarily mutually exclusive for the present findings. First, imaging studies were not the only method of cancer detection used. Other modalities, such as physical examination and patient symptoms, could have complemented radiologic investigations and compensated for the hypothetically less sensitive regular-interval follow-up. Another possible explanation is the slow progression of "occult" breast cancers. Therefore, the cancer did not appreciably differ, in terms of the current TNM cancer staging systems, when

Table III. Comparison of known or unknown cancer status at 24 months for women with BI-RADS 3 screen images.

Characteristic	No (%)		p-value[a]
	Known cancer status (n = 846)	Unknown cancer status (n = 190)	
Mean age ± SD (years)	49.9 ± 6.9	49.8 ± 8.1	0.839
Dense breast	771 (91.1)	167 (87.9)	0.168
Biopsy	36 (4.3)	2 (1.1)	0.034
Interval progression[b]	29 (3.4)	4 (2.1)	0.348
Type of follow-up			
Short-interval	278 (32.9)	90 (47.4)	< 0.001
Regular-interval	568 (67.1)	100 (52.6)	-

[a] This is the p-value according to t-test or chi-square test as appropriate.
[b] The interval progression within 12 months of follow-up.
BI-RADS: Breast Imaging and Data Reporting System; SD: standard deviation

detected by either short- or regular-interval follow-up images.[6,10,12] In addition, commonly used imaging studies (mammography and ultrasonography [US]) might not be sensitive to changes in the size of initially small tumours.[6] Finally, the lack of any significant differences could have occurred by chance (see below).

The 15% relative frequency of BI-RADS 3 images in the present study was considerably higher than the frequency noted in previous studies (1.1 to 12.2%).[13,14] However, the 0.7% PPV in the present study was comparable to the accepted risk for women with BI-RADS 3 lesions.[15] The use of the US as well as mammography in determining the BI-RADS category might partly explain the apparently high relative frequency of BI-RADS 3 in the present study, by including more women with benign lesions into the BI-RADS 3 category.[6]

There is some evidence that the addition of US to mammography can significantly increase the rate of detection of early breast cancer, although with a concurrent increase in the rate of false positive findings.[15,16] US is most useful for younger women or women with dense breasts.[15,16] Since the combined use of US and mammography is standard practice in our institution, as can be seen by the high proportion of women with dense breasts (86%) (Table I), this study has not looked at the effects of US and mammography separately.

Some limitations of the present study include its non-randomised design, the small number of outcomes and incomplete data. Because of the small number of outcomes, the lack of evidence for any significant differences could have occurred purely by chance. A relatively high percentage of women who did not complete the 24-month follow-up (18%, 190 out of 1,036) might have introduced some selection bias into the comparison between the two groups. In particular,

significantly more women with unknown cancer status underwent short-interval follow-up than those with known cancer status (Table III). If women with unknown cancer status are more likely to have breast cancer, then the results of the present study might have been biased toward a lower cancer detection rate in the short-interval follow-up group.

Even if it is true that short-interval follow-up for asymptomatic women with BI-RADS 3 breast images does not in general increase the rate of image-detected breast cancer or detect breast cancers at an earlier stage, there may be subgroups of higher-risk women with rapidly progressive breast cancer who may benefit from short-interval follow-up, or may even require a biopsy. Methods to identify such women include computer-aided classification (CAC),[17] statistical models,[18,19] the use of longitudinal imaging information,[20] and the concurrent use of other imaging studies.[21,22] The least expensive method would be to use easily obtainable clinical or radiologic information to quantify the breast cancer risk based on statistical models. For a more valid and reliable comparison of the relative effectiveness of different follow-up intervals, a larger randomised study needs to be conducted in the future.

Since higher-risk women with category 3 breast images cannot be reliably identified with the present technology, the authors recommend a middle ground, that has been previously suggested,[6] between not performing short-interval follow-up at all and obtaining short-interval follow-up for two to three years. In other words, it is recommended that short-interval follow-up be conducted only once, and if no interval progression is noted, the patient should be followed up annually thereafter.

This study found no significant differences in terms of breast image stability, the cancer detection rate and the

cancer stage at detection between the short- and regular-interval imaging follow-up strategies in women with BI-RADS 3 screen images. The savings associated with the decrease in the number of imaging studies when using the regular follow-up strategy, as well as the reduction in the "induced" costs (including additional imaging studies and biopsy procedures), may be substantial, especially in the setting of a developing economy.

REFERENCES

1. Orel SG, Kay N, Reynolds C, Sullivan DC. BI-RADS categorization as a predictor of malignancy. Radiology 1999; 211:845-50.
2. Yasmeen S, Romano PS, Pettinger M, et al. Frequency and predictive value of a mammographic recommendation for short-interval follow-up. J Natl Cancer Inst 2003; 95:429-36.
3. 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; 222:529-35.
4. American College of Radiology. Breast imaging reporting and data system (BI-RADS). 4th ed. Reston: American College of Radiology, 2003.
5. Rubin E. Six-month follow-up: an alternative view. Radiology 1999; 213:15-8.
6. Varas X, Leborgne JH, Leborgne F, et al. Revisiting the mammographic follow-up of BI-RADS category 3 lesions. AJR Am J Roentgenol 2002; 179:691-5.
7. Gokalp G, Topal U. MR imaging in probably benign lesions (BI-RADS category 3) of the breast. Eur J Radiol 2006; 57:436-44.
8. 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.
9. Kerlikowske K, Smith-Bindman R, Abraham LA, et al. Breast cancer yield for screening mammographic examinations with recommendation for short-interval follow-up. Radiology 2005; 234:684-92.
10. Aiellow Bowles EJ, Miglioretti DL, Sickles EA, et al. Accuracy of short-interval follow-up mammograms by patient and radiologist characteristics. AJR Am J Roentgenol 2008; 190:1200-8.
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. Vizcaino I, Gadea L, Andreo L, et al. Short-term follow-up results in 795 nonpalpable probably benign lesions detected at screening mammography. Radiology 2001; 219:475-83.
13. Monticciolo DL, Caplan LS. The American College of Radiology's BI-RADS 3 Classification in a Nationwide Screening Program: current assessment and comparison with earlier use. Breast J 2004; 10:106-10.
14. Kerlikowske K, Smith-Bindman R, Ljung BM, Grady D. Evaluation of abnormal mammography results and palpable breast abnormalities. Ann Intern Med 2003; 139:274-84.
15. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. Radiology 2002; 225:165-75.
16. Berg WA, Blume JD, Cormack JB, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. JAMA 2008; 299:2151-63.
17. Buchbinder SS, Leichter IS, Lederman RB, et al. Computer-aided classification of BI-RADS category 3 breast lesions. Radiology 2004; 230:820-3.
18. Burnside ES, Rubin DL, Fine JP, et al. Bayesian network to predict breast cancer risk of mammographic microcalcifications and reduce number of benign biopsy results: initial experience. Radiology 2006; 240:666-73.
19. Vacek PM, Geller BM. A prospective study of breast cancer risk using routine mammographic breast density measurements. Cancer Epidemiol Biomarkers Prev 2004; 13:715-22.
20. Kerlikowske K, Ichikawa L, Miglioretti DL, et al. Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. J Natl Cancer Inst 2007; 99:386-95.
21. Lord SJ, Lei W, Craft P, et al. A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer. Eur J Cancer 2007; 43:1905-17.
22. Warner E, Messersmith H, Causer P, et al. Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. Ann Intern Med 2008; 148:671-9.