Published in final edited form as:

AJR Am J Roentgenol. 2010 April ; 194(4): 1152–1159. doi:10.2214/AJR.09.3064.

Recommendation for Short-Interval Follow-Up Examinations After a Probably Benign Assessment: Is Clinical Practice Consistent With BI-RADS Guidance?

Erin J. Aiello Bowles¹, Edward A. Sickles², Diana L. Miglioretti¹, Patricia A. Carney³, and Joann G. Elmore⁴

¹Group Health Research Institute, 1730 Minor Ave., Ste. 1600, Seattle, WA 98101

²Department of Radiology, University of California, San Francisco, San Francisco, CA

³Department of Family Medicine, Oregon Health & Science University, Portland, OR

⁴Department of Internal Medicine, University of Washington School of Medicine, Seattle, WA

Abstract

Objective—American College of Radiology BI-RADS guidance suggests that women with a probably benign finding on mammography receive a management recommendation for short-interval follow-up; historically, radiologists in community practice have not consistently linked this assessment with short-interval follow-up. We evaluated predictors of discordance between probably benign assessments and short-interval follow-up recommendations.

Materials and Methods—We linked data on 196 radiologists who completed a survey on demographic and practice patterns to 15,515 diagnostic mammograms they interpreted with probably benign assessments between 2001 and 2006. Patient characteristics were collected at the time of the mammography. Using logistic regression, we examined whether patient and radiologist characteristics were associated with the odds of short-interval follow-up recommendations (relative to a recommendation for normal follow-up, additional imaging evaluation, or biopsy or surgical consultation).

Results—Overall, 90.9% of mammograms with probably benign findings were recommended for short-interval follow-up; 4.3% were recommended for normal follow-up, 3.0% for additional imaging, and 1.8% for biopsy or surgical consultation. Women with probably benign findings were less likely to receive a short-interval follow-up recommendation if they had extremely dense breasts versus almost entirely fatty breasts (odds ratio [OR], 0.61; 95% CI, 0.39–0.96) or had a breast lump versus no symptoms (OR, 0.55; 95% CI, 0.38–0.78). Radiologists were less likely to recommend short-interval follow-up if they had ≥ 20 years of experience versus < 10 years of experience (OR, 0.57; 95% CI, 0.36–0.90) but more likely if they practiced primarily at an academic medical center versus other institutions (OR, 2.66; 95% CI, 1.14–6.21).

Conclusion—In contrast to older studies, the majority of probably benign assessments are now recommended for short-interval follow-up, but the probability of short-interval follow-up recommendations varies by patient and radiologist characteristics.

Keywords

breast cancer; diagnostic mammography; guidelines; radiologists; short-interval follow-up mammography

BI-RADS, which was developed by the American College of Radiology, outlines a standardized reporting system of mammographic lesions. The *BI-RADS Atlas* [1] outlines seven assessment categories, each of which has an associated follow-up recommendation to guide radiologists and referring clinicians in the management of their patients [2]. BI-RADS assessment category 3 (probably benign finding) is associated with a suggested recommendation for short-interval (< 1 year, usually 6 months) follow-up. The lesions intended for a category 3 assessment are those believed to have an extremely low (< 2%) probability of being malignant [1,3–9]. Short-interval follow-up mammography monitors lesions for changes at a more frequent interval than regular screening and is intended to serve as an alternative to invasive procedures, such as biopsy or fine-needle aspiration [1,10]. Given the low probability of malignancy, discordance between probably benign assessments and short-interval follow-up recommendation can lead to unnecessary and costly workup for the majority of findings that are truly benign.

Previous studies have reported that up to 14% of screening and diagnostic mammography examinations are assigned BI-RADS category 3 [4–6,11–14]. These studies have also shown that only 40–71% of these examinations are given the suggested recommendation for short-interval follow-up [4,11,13,15]. The inconsistency between BI-RADS category 3 assessment and short-interval follow-up recommendation may be related to patient characteristics, such as age, family history of breast cancer, body mass index, and breast density [5,12,14,15], or to variation in the practice patterns of individual radiologists [12,14]. Some of the studies describing the use of short-interval follow-up recommendations took place before 1999, before BI-RADS was used as widely in clinical practice as it is now [5]. Fortunately, concordance has increased over time (from 54% between 1996 and 1999 to 71% between 1999 and 2001), perhaps reflecting changes in the *BI-RADS Atlas* [1] and Mammography Quality Standards Act [13]. However, no study has evaluated concordance since publication of the most recent edition of *BI-RADS Atlas* in 2003 [1], the only edition that provides more explicit guidance concerning the linkage of assessment categories and management recommendations.

Using data from the Breast Cancer Surveillance Consortium, we assessed current trends in the use of short-interval follow-up recommendations for probably benign assessments and examined whether concordant recommendations are associated with patient or radiologist characteristics. A recommendation for prompt additional imaging, biopsy, or surgical consultation instead of short-interval follow-up mammography may lead to unnecessary patient anxiety, workup, and increased health care costs. A recommendation for normal follow-up instead of short-interval follow-up may lead to delay in diagnosis of the infrequent (< 2%) cancers found among probably benign lesions. Understanding the factors that influence a radiologist's management of probably benign lesions may help in designing practice-based interventions to reduce variability in mammographic performance.

Materials and Methods

Study Population

This analysis used data from seven mammography registries that are part of the National Cancer Institute–funded Breast Cancer Surveillance Consortium [16]: Group Health, Western Washington; Colorado Mammography Project; New Hampshire Mammography Network; Carolina Mammography Registry, North Carolina; New Mexico Mammography Project; San Francisco Mammography Registry, California; and Vermont Breast Cancer Surveillance System. The Breast Cancer Surveillance Consortium statistical coordinating center pooled data for analysis. Each registry and the statistical coordinating center received institutional review board approval to enroll participants, link data, and perform analytic studies. All procedures were HIPAA compliant, and all registries and the statistical coordinating center have received

a federal certificate of confidentiality and other protection for the identities of women, physicians, and facilities who are subjects of this research.

Data Collection

Radiologist characteristics—We invited radiologists who interpreted mammograms during 2005–2006 at any of the Breast Cancer Surveillance Consortium sites to complete a self-administered mailed survey between January 2006 and September 2007. Surveys were distributed in four sites (Colorado, North Carolina, New Hampshire, and Washington) in 2006 and in the remaining three sites (California, New Mexico, and Vermont) in 2007. We linked survey results from 196 participating radiologists to the diagnostic mammograms they interpreted between 2001 and 2006. A copy of the survey is available at www.breastscreening.cancer.gov/collaborations/favor.html.

Patient characteristics—The Breast Cancer Surveillance Consortium collected information on patient demographics (e.g., age and race) and breast cancer risk factors (e.g., family history of breast cancer, breast symptoms, menopausal status, and current hormone therapy use) using standardized questionnaires completed by women at the time of their mammographic examinations [16].

BI-RADS category 3 assessments—We included diagnostic mammography examinations performed for the additional evaluation of a recent mammogram and given a final BI-RADS assessment of category 3 (probably benign) during 2001–2006. We did not include diagnostic examinations performed to evaluate a breast problem, an indication usually used for symptomatic women. This classification was based on the radiologist's indication for the examination; however, women could still self-report having symptoms or breast problems at the time of the examination regardless of the radiologist's indication. *BI-RADS Atlas* [1] recommends that radiologists give probably benign assessments for 2 to 3 years after the first probably benign assessment to evaluate stability versus change in the lesion. However, only the first two follow-up examinations in the surveillance protocol are recommended at an interval shorter than routine annual screening mammography (usually 6 months). If women had multiple examinations with a probably benign assessment, we included their first mammography with a probably benign assessment. We excluded mammography examinations for women with a history of breast cancer or a prior short-interval follow-up mammography. Finally, we excluded mammograms that were missing key covariate or outcome information, including breast density, time since previous mammography, and the final follow-up recommendation made by the radiologist. The final sample included 15,515 women, which was 5.6% of all women from the same time period without missing data who underwent mammography indicated by the radiologist to be for the additional evaluation of a recent mammogram ($n = 277,446$).

Mammography recommendations—Recommendations were classified as short-interval follow-up, normal follow-up, additional imaging evaluation, and biopsy or surgical evaluation or clinical examination from the radiologists' reports. All short-interval follow-up recommendations were included regardless of the recommended time interval for follow-up, although 84% of the 15,515 examinations were recommended for follow-up in 6 months. Recommendations for additional imaging evaluation in 6 months ($n = 640$) were classified as short-interval follow-up instead of additional imaging evaluation. In a sensitivity analysis, keeping the original recommendation for additional imaging evaluation did not change our results. Some radiology facilities use software that automatically links assessment and recommendation based on BI-RADS guidelines; however, in some systems, it is possible to override the linkage and make a different recommendation. For each mammography, we

included an indicator of whether the mammographic recommendation was linked to the final assessment at the time of interpretation as a covariate in the analysis.

Statistical Analysis

We describe the characteristics of women who received probably benign assessments and of the radiologists who made the probably benign assessments. We also describe the distribution of patient and radiologist characteristics by follow-up recommendation after the probably benign assessment.

We used logistic regression to examine the association between the odds of receiving a recommendation for a short-interval follow-up mammography given a probably benign assessment and patient and radiologist characteristics. Using generalized estimating equations [17], we account for clustering by radiologist and adjusted for patient age, year of the mammography, mammographic breast density, and Breast Cancer Surveillance Consortium registry.

Results

In Table 1, we show the distribution of radiologist characteristics among those who gave BI-RADS category 3 assessments. About 70% of radiologists were male, and 65% had more than 10 years of experience interpreting mammograms. Few radiologists spent the majority of their time in breast imaging; 84.7% spent less than 25% of their workload interpreting screening mammograms, and 72.1% spent less than 10% of their workload interpreting diagnostic mammograms.

Among 15,515 women with probably benign assessments, 64.6% were between the ages of 40 and 59 years, and 16.1% reported having a first-degree family history of breast cancer (Table 2). The majority of examinations with probably benign assessments occurred among women who reported no symptoms at the time of mammography. Even though we only included examinations that were indicated by the radiologist to be for additional evaluation, 1,159 women reported having breast symptoms (a lump or other symptoms) at the time of their examination. Mammographic assessment and recommendation were linked among 41.4% of mammographic examinations.

Table 2 also shows the distribution of follow-up recommendations by patient characteristics. Almost 91% of women with a probably benign assessment received a recommendation for a short-interval follow-up mammography, but 4.3% were recommended for normal follow-up, 3.0% for additional imaging, and 1.8% for biopsy or surgical or clinical evaluation. Women were less likely to be recommended for short-interval follow-up if they were postmenopausal (odds ratio [OR], 0.78; 95% CI, 0.61–0.98), had extremely dense breasts (OR, 0.61; 95% CI, 0.38–0.96), or reported a breast lump (OR, 0.55; 95% CI, 0.38–0.78). A higher proportion of women with extremely dense breasts tended to be recommended for more aggressive management in the form of additional imaging or biopsy, compared with women with fatty breasts. Women with symptoms were also more likely to be recommended for breast biopsy, surgical consultation, or clinical evaluation than were women without symptoms. The adjusted odds of a short-interval follow-up recommendation were statistically significantly increased from 2003 through 2006 compared with 2001 and for women who had mammographic examinations at facilities that linked assessment with recommendation (OR, 18.05; 95% CI, 7.8–41.77). The increase in short-interval follow-up recommendations over time and among facilities with linked assessments and recommendations was accompanied by decreases in all other types of recommendations.

Radiologists who were ≥ 55 years old or who had more than 20 years of experience interpreting mammograms were less likely to give a short-interval follow-up recommendation (OR, 0.48; 95% CI, 0.26–0.88) than were younger radiologists and radiologists with fewer years of experience interpreting mammograms (OR, 0.53; 95% CI, 0.29–0.95) (Table 3). Radiologists with a primary appointment at an academic medical center were statistically significantly more likely to recommend short-interval follow-up (OR, 2.66; 95% CI, 1.14–6.21). In addition, radiologists with a total mammography volume $\geq 1,000$ annually were more likely to recommend short-interval follow-up compared with radiologists with an annual volume $< 1,000$.

Discussion

We found that most diagnostic mammograms assessed as probably benign were given a recommendation for short-interval follow-up as suggested by the *BI-RADS Atlas* [1] and that the percentage of short-interval follow-up recommendations increased over time, reaching 96.7% in 2006. These are much higher proportions of short-interval follow-up recommendations after BI-RADS category 3 assessments than reported in previous studies [4,11,13,15]. One explanation for the increased consistency between assessment and recommendation in our study is that we limited our analysis to diagnostic examinations that were performed for the additional evaluation of a recent mammography. Previous studies showing lower consistency included screening mammographic examinations, examinations for which BI-RADS category 3 assessments should not be made at all [4,11,13,15]. Only Geller et al. [13] evaluated concordance rates among diagnostic examinations that were performed to evaluate a breast problem, but even these rates of concordance were lower than those we found (43.8% in 1996 vs 55.8% in 2001). We did not include diagnostic examinations performed to evaluate a breast problem in our study, which also could have accounted for some of the difference in results. Finally, our finding of increased consistency over time may be the result of additional guidance put forth by the American College of Radiology in the fourth edition of BI-RADS in 2004 [1].

Mammographic examinations that were interpreted using a system that automatically linked assessment and recommendation were associated with the highest odds of short-interval follow-up recommendations. Over 99% of these mammographic examinations received short-interval follow-up recommendations, though a small proportion (0.9%) were given different recommendations, suggesting that radiologists infrequently overrode the linked short-interval follow-up recommendation. This result is consistent with those of a previous study by Geller et al. [18], who showed that linked systems improved the concordance between BI-RADS assessments and recommendations. Approximately 15% of mammograms in the study by Geller et al. were interpreted using linked systems, compared with 40% in the present article. Systems that automatically link assessment and recommendation are increasingly being used in practice, and linkage may be the most effective way to improve agreement, not only among BI-RADS category 3 assessments but also among all BI-RADS assessment categories.

Use of short-interval follow-up recommendations for lesions assessed as probably benign varied by both patient and radiologist characteristics. Both increased mammographic breast density and reporting of breast symptoms at the time of mammography were associated with a decreased probability of receiving a short-interval follow-up recommendation, consistent with the probably benign assessment. Instead, these women tended to receive recommendations for immediate follow-up (additional imaging or biopsy). It is important to note that few of the 15,515 probably benign assessments were given to women with extremely dense breasts ($n = 1,005$) or with symptoms ($n = 266$ with a breast lump and $n = 893$ with other symptoms). Radiologists may have recommended these women for immediate follow-up because they doubted the accuracy of mammography in women with dense breasts [19,20] or because they

considered symptomatic women to be at greater than average breast cancer risk [21]. The majority of mammography malpractice suits citing delays in cancer diagnoses are initiated by women with symptoms, which could help explain the more aggressive recommendations for this group [22].

Radiologist characteristics were also associated with the consistency between BI-RADS category 3 assessments and short-interval follow-up recommendations. Older radiologists (≥ 55 years old) and those with more than 20 years of experience were less likely to recommend short-interval follow-up compared with younger radiologists or radiologists with fewer years of experience. These two characteristics are highly correlated and may suggest that radiologists who have been practicing longer are more resistant to changing their ways of practice when new or revised interpretive guidance is published. In addition, these radiologists were trained long before the revised American College of Radiology guidance was published, suggesting that there is a need for continuing education for all radiologists, regardless of age or years of experience. Several studies have shown that changing established practice patterns among physicians is difficult and often requires multiple interventions [23,24]. Training in BI-RADS specifically has been shown to improve agreement in final mammogram assessments among radiologists [25], but we are unaware of any studies that have evaluated whether BI-RADS training improves agreement between recommendations and assessments.

Radiologists who were more likely to recommend short-interval follow-up after a probably benign examination included those with a primary affiliation at an academic medical center. These radiologists may be more up to date on BI-RADS guidance or may have more opportunities for continuing education in mammography than do radiologists who are not affiliated with an academic medical center. In addition, radiologists who reported interpreting $\geq 1,000$ mammograms in a single year were more likely to recommend short-interval follow-up, compared with radiologists who reported interpreting $< 1,000$ mammograms. We are not aware of other studies that have evaluated radiologist characteristics, such as age, experience, academic affiliation, or volume in relation to the consistency of following the BI-RADS guidance of linking mammography assessments and recommendations.

One limitation of our study was that we evaluated radiologist characteristics from a survey collected at one point in time with mammography assessments and recommendations that changed over time. To account for this fact, we recoded radiologist age and experience so that they changed over time with the year of the mammography interpretation. However, most other radiologist characteristics, such as sex and academic affiliation, do not change over time. In addition, our study population was limited to women with their first diagnostic examination assessed as probably benign and cannot be generalized to subsequent examinations with probably benign assessments. Our data may be subject to some misclassification if women had prior mammographic examinations outside of the Breast Cancer Surveillance Consortium registries that we could not identify. Finally, we could not assess patient preferences for follow-up. Some women may have requested an immediate biopsy or additional imaging instead of waiting 6 months for another mammography, which could have accounted for some of the variability in recommendations.

Our study also has several strengths. We were able to link a large number of mammographic examinations with both patient characteristics and characteristics of the radiologists who interpreted them. Our study population included diagnostic mammographic examinations performed for the additional evaluation of a recent screening examination across seven U.S. states, in Breast Cancer Surveillance Consortium sites that collect data from facilities that range from small rural practices to large urban hospitals. Therefore, our results should be representative of community practice in the United States [3].

In conclusion, we noted a high degree of consistency between radiologists' recommendations for short-interval follow-up after a probably benign assessment and BI-RADS guidance; this consistency improved between 2001 and 2006. The use of computerized systems that automatically link assessment and recommendations may be an effective method for further increasing consistency with BI-RADS guidance. In addition, continuing medical education on BI-RADS guidance would be relevant for all practicing mammographers to reduce variability in assessment and recommendation and, ultimately, to improve mammographic performance.

Acknowledgments

The authors had full responsibility in the design of the study, the collection of the data, the analysis and interpretation of the data, the decision to submit the manuscript for publication, and the writing of the manuscript. We thank the participating women, mammography facilities, and radiologists for the data they have provided for this study. A list of the Breast Cancer Surveillance Consortium investigators and procedures for requesting Breast Cancer Surveillance Consortium data for research purposes are provided at breastscreening.cancer.gov/.

This work was supported by the National Cancer Institute (grants 1R01 CA107623 and 1K05 CA104699; and grants U01CA63740, U01CA86076, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, and U01CA70040 to the Breast Cancer Surveillance Consortium), the Agency for Healthcare Research and Quality (grant 1R01 CA107623), the Breast Cancer Stamp Fund, and the American Cancer Society, through a generous donation from the Longaberger Company's Horizon of Hope Campaign (grants SIRGS-07-271-01, SIRGS-07-272-01, SIRGS-07-273-01, SIRGS-07-274-01, SIRGS-07-275-01, and SIRGS-06-281-01). The collection of cancer data used in this study was supported in part by several state public health departments and cancer registries throughout the United States. A full description of these sources is available at breastscreening.cancer.gov/work/acknowledgement.html.

References

1. American College of Radiology. Breast imaging reporting and data system: BI-RADS atlas. 4th. Reston, VA: American College of Radiology; 2003. BI-RADS: mammography.
2. D'Orsi CJ. The American College of Radiology mammography lexicon: an initial attempt to standardize terminology. *AJR* 1996;166:779–780. [PubMed: 8610548]
3. Sickles EA, Miglioretti DL, Ballard-Barbash R, et al. Performance benchmarks for diagnostic mammography. *Radiology* 2005;235:775–790. [PubMed: 15914475]
4. 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;234:684–692. [PubMed: 15734926]
5. 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;95:429–436. [PubMed: 12644536]
6. Varas X, Leborgne JH, Leborgne F, Mezzera J, Jaumandreu S, Leborgne F. Revisiting the mammographic follow-up of BI-RADS category 3 lesions. *AJR* 2002;179:691–695. [PubMed: 12185047]
7. Sickles EA. Periodic mammographic follow-up of probably benign lesions: results in 3,184 consecutive cases. *Radiology* 1991;179:463–468. [PubMed: 2014293]
8. 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;219:475–483. [PubMed: 11323475]
9. Varas X, Leborgne F, Leborgne JH. Nonpalpable, probably benign lesions: role of follow-up mammography. *Radiology* 1992;184:409–414. [PubMed: 1620838]
10. Aiello Bowles EJ, Miglioretti DL, Sickles EA, et al. Accuracy of short-interval follow-up mammograms by patient and radiologist characteristics. *AJR* 2008;190:1200–1208. [PubMed: 18430832]
11. 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. [PubMed: 11818624]

12. Caplan LS, Blackman D, Nadel M, Monticciolo DL. Coding mammograms using the classification “probably benign finding—short interval follow-up suggested”. *AJR* 1999;172:339–342. [PubMed: 9930778]
13. 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. [PubMed: 11818625]
14. 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;10:106–110. [PubMed: 15009036]
15. Lehman C, Holt S, Peacock S, White E, Urban N. Use of the American College of Radiology BI-RADS guidelines by community radiologists: concordance of assessments and recommendations assigned to screening mammograms. *AJR* 2002;179:15–20. [PubMed: 12076896]
16. Ballard-Barbash R, Taplin SH, Yankaskas BC, et al. Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database. *AJR* 1997;169:1001–1008. [PubMed: 9308451]
17. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986;42:121–130. [PubMed: 3719049]
18. Geller BM, Ichikawa LE, Buist DS, et al. Improving the concordance of mammography assessment and management recommendations. *Radiology* 2006;241:67–75. [PubMed: 16990672]
19. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 2000;92:1081–1087. [PubMed: 10880551]
20. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007;356:227–236. [PubMed: 17229950]
21. Egger JR, Cutter GR, Carney PA, et al. Mammographers' perception of women's breast cancer risk. *Med Decis Making* 2005;25:283–289. [PubMed: 15951455]
22. Mitnick JS, Vazquez MF, Kronovet SZ, Roses DF. Malpractice litigation involving patients with carcinoma of the breast. *J Am Coll Surg* 1995;181:315–321. [PubMed: 7551325]
23. Satterlee WG, Eggers RG, Grimes DA. Effective medical education: insights from the Cochrane Library. *Obstet Gynecol Surv* 2008;63:329–333. [PubMed: 18419832]
24. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance: a systematic review of the effect of continuing medical education strategies. *JAMA* 1995;274:700–705. [PubMed: 7650822]
25. 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. [PubMed: 12202727]

TABLE 1
Characteristics of 196 Radiologists Who Gave BI-RADS 3 Assessments on Diagnostic Mammography Examinations for the Additional Evaluation of a Recent Mammography Between 2001 and 2006

Characteristic	No. (%) ^a of Radiologists (<i>n</i> = 196)
Age (y) ^b	
< 44	66 (33.7)
45–54	75 (38.3)
≥ 55	55 (28.1)
Sex	
Male	138 (70.4)
Female	58 (29.6)
Academic affiliation	
No affiliation	155 (79.9)
Adjunct, appointment	17 (8.8)
Primary, appointment	22 (11.3)
Missing	2
Years interpreting mammograms ^d	
< 10	67 (34.5)
10–19	74 (38.1)
≥ 20	53 (27.3)
Missing	2
Hours per week spent in breast imaging	
≤ 20	118 (63.4)
21–40	38 (20.4)
> 40	30 (16.1)
Missing	10
Hours worked per week	
≤ 20	13 (6.7)
21–40	63 (32.6)
> 40	117 (60.6)
Missing	3
Total estimated mammogram volume in 2005	
≤ 999	18 (9.9)
1,000–1,999	57 (31.3)
2,000–4,999	77 (42.3)
≥ 5,000	30 (16.5)
Missing	14
Self-reported screening volume in 2005	
≤ 999	36 (19.4)
1,000–1,999	56 (30.1)
2,000–4,999	74 (39.8)

Characteristic	No. (%) ^a of Radiologists (<i>n</i> = 196)
≥ 5,000	20 (10.8)
Missing	10
Self-reported diagnostic volume in 2005	
< 250	67 (36.4)
250–499	53 (28.8)
500–999	37 (20.1)
≥ 1,000	27 (14.7)
Missing	12
Portion of workload interpreting screening mammograms (%)	
0–10	76 (40.0)
11–25	85 (44.7)
> 25	29 (15.3)
Missing	6
Portion of workload interpreting diagnostic mammograms (%)	
0–10	137 (72.1)
11–20	22 (11.6)
21–30	14 (7.4)
31–40	7 (3.7)
> 40	10 (5.3)
Missing	6

^a Percentages were calculated excluding missing data.

^b Age and years of interpretation were allowed to vary with year of examination.

TABLE 2
Association Between Patient Characteristics and Final Follow-Up Recommendation Among Mammographic Examinations With a BI-RADS 3 Assessment

Characteristic	No. of Examinations	Overall % ^a	Normal Follow-Up (%) (<i>n</i> = 660)	Short-Interval Follow-Up (%) (<i>n</i> = 13,967)	Additional Imaging (%) (<i>n</i> = 468)	Biopsy, Surgical Evaluation, Clinical Examination (%) (<i>n</i> = 266)	Adjusted OR ^b (95% CI), Short-Interval Follow-Up Recommendation vs Other
Overall	15,515		4.3	90.9	3.0	1.8	
Age (y)							
< 40	712	4.6	3.5	88.1	3.8	4.6	0.88 (0.66–1.17)
40–49	5,277	34.0	3.8	91.7	2.9	1.6	Reference
50–59	4,747	30.6	3.9	91.2	3.3	1.6	0.92 (0.76–1.11)
60–69	2,631	17.0	5.0	90.5	2.7	1.7	0.97 (0.76–1.26)
≥ 70	2,148	13.8	5.7	89.9	2.9	1.5	0.87 (0.69–1.11)
Menopausal status							
Premenopausal	4,903	36.1	3.4	91.4	3.3	1.9	Reference
Perimenopausal	382	2.8	1.8	94.8	1.6	1.8	0.65 (0.37–1.15)
Postmenopausal	8,284	61.1	5.0	90.2	3.0	1.7	0.78 (0.61–0.98)
Missing	1,946						
Hormone therapy use							
No	6,575	81.5	7.4	84.9	5.2	2.6	Reference
Yes	1,493	18.5	9.0	82.3	6.0	2.7	0.85 (0.69–1.04)
Missing	7,447						
Breast density							
Almost entirely fat	647	4.2	6.8	90.4	2.2	0.6	Reference
Scattered fibroglandular tissue	6,418	41.4	4.2	90.3	3.8	1.7	1.06 (0.72–1.56)
Heterogeneously dense	7,445	48.0	4.1	91.8	2.4	1.7	0.94 (0.62–1.43)
Extremely dense	1,005	6.5	4.8	88.7	3.8	2.8	0.61 (0.38–0.96)
Family history							
No	12,023	83.9	4.5	90.3	3.3	1.9	Reference
Yes	2,300	16.1	4.3	91.7	2.3	1.6	0.97 (0.8–1.17)
Missing	1,192						
Breast symptoms							

Characteristic	No. of Examinations	Overall % ^a	Normal Follow-Up (%) (n = 660)	Short-Interval Follow-Up (%) (n = 13,967)	Additional Imaging (%) (n = 468)	Biopsy, Surgical Evaluation, Clinical Examination (%) (n = 266)	Adjusted OR ^b (95% CI), Short-Interval Follow-Up Recommendation vs Other
None	6,287	84.4	8.4	83.6	5.3	2.6	Reference
Other	893	12.0	8.0	80.7	9.2	2.1	0.55 (0.27–1.15)
Lump	266	3.6	6.8	75.6	7.1	10.5	0.55 (0.38–0.78)
Missing	8,069						
Race							
White	12,476	84.4	4.0	91.4	2.9	1.7	Reference
Black	692	4.7	11.0	81.9	5.8	1.3	1.28 (0.97–1.71)
Asian	398	2.7	1.0	94.5	1.5	3.0	0.95 (0.61–1.47)
Other	1,214	8.2	3.9	90.4	3.1	2.6	1.16 (0.75–1.80)
Missing	735						
Year of examination							
2001	3,583	23.1	5.1	88.5	4.3	2.0	Reference
2002	2,995	19.3	5.6	87.9	3.9	2.5	1.09 (0.82–1.46)
2003	2,960	19.1	4.0	91.7	2.4	2.0	1.62 (1.14–2.32)
2004	2,467	15.9	4.2	92.0	2.8	1.0	2.01 (1.26–3.19)
2005	1,895	12.2	3.4	92.7	2.5	1.4	1.84 (1.14–2.97)
2006	1,615	10.4	1.7	96.7	0.9	0.8	2.33 (1.35–4.02)
Linked assessment and recommendation							
No	9,095	58.6	7.1	85.1	5.1	2.7	Reference
Yes	6,420	41.4	0.3	99.2	0.1	0.5	18.05 (7.8–41.77)

^a Percentages in this column were calculated excluding missing values.

^b Odds ratios (ORs) were adjusted for multiple examinations per radiologist, Breast Cancer Surveillance Consortium registry, age, density, and time since previous mammography. Bold ORs are statistically significant compared with the reference group.

TABLE 3
Association Between Radiologist Characteristics and Recommendation Given After Assigning a BI-RADS Category 3 Assessment to a Mammogram

Characteristic	All Examinations (No.) ^a	Normal Follow-Up (%) (n = 660)	Short-Interval Follow-Up, (%) (n = 13,967)	Additional Imaging (%) (n = 468)	Biopsy, Surgical Evaluation, Clinical Examination (%) (n = 266)	Adjusted OR ^b (95% CI) for Short-Interval Follow-Up Recommendation vs Other
Overall	15,515	4.3	90.9	3.0	1.8	
Age (y) ^c						
< 44	5,383	5.1	90.8	2.6	1.5	Reference
45–54	5,583	3.6	92.1	2.7	1.6	0.63 (0.38–1.03)
≥ 55	4,549	4.2	89.6	3.9	2.2	0.48 (0.26–0.88)
Sex						
Male	9,460	5.9	87.7	4.0	2.4	Reference
Female	6,055	1.8	96.0	1.5	0.8	1.73 (0.83–3.59)
Academic affiliation						
No affiliation	11,197	5.3	88.8	3.7	2.1	Reference
Adjunct, appointment	1,176	3.2	95.0	0.7	1.1	1.53 (0.63–3.68)
Primary, appointment	3,113	1.1	96.9	1.4	0.6	2.66 (1.14–6.21)
Years interpreting mammograms ^b						
< 10	5,042	4.2	92.1	2.5	1.2	Reference
10–19	6,798	3.1	92.3	2.8	1.8	0.65 (0.41–1.03)
≥ 20	3,652	6.5	86.8	4.2	2.5	0.53 (0.29–0.95)
Hours per week spent in breast imaging						
≤ 20	8,322	4.7	89.8	3.5	2.0	Reference
21–40	3,718	3.7	91.5	3.0	1.8	1.17 (0.64–2.13)
> 40	2,550	3.7	93.0	2.4	1.0	1.53 (0.67–3.48)
Hours worked per week						
≤ 20	270	9.6	78.9	8.1	3.3	Reference
21–40	6,725	2.8	93.1	2.1	2.0	1.97 (0.72–5.39)
> 40	8,038	5.6	89.1	3.8	1.5	2.21 (0.85–5.80)
Total estimated mammogram volume in 2005						

Characteristic	All Examinations (No.) ^a	Normal Follow-Up (%) (n = 660)	Short-Interval Follow-Up, (%) (n = 13,967)	Additional Imaging (%) (n = 468)	Biopsy, Surgical Evaluation, Clinical Examination (%) (n = 266)	Adjusted OR ^b (95% CI) for Short-Interval Follow-Up Recommendation vs Other
Self-reported screening volume in 2005						Reference
≤ 999	589	7.5	80.3	9.3	2.9	2.97 (1.54–5.70)
1,000–1,999	3,468	4.5	91.4	2.7	1.5	2.32 (1.13–4.77)
2,000–4,999	7,030	4.4	90.0	3.2	2.4	3.14 (1.46–6.76)
≥ 5,000	3,469	3.5	92.9	2.7	0.9	
Self-reported diagnostic volume in 2005						Reference
≤ 999	2,523	5.6	89.0	3.8	1.6	1.69 (0.83–3.41)
1,000–1,999	2,661	3.3	91.5	3.0	2.2	1.27 (0.76–2.12)
2,000–4,999	7,801	4.0	91.2	2.9	1.9	1.55 (0.75–3.19)
≥ 5,000	1,899	5.4	90.2	3.3	1.1	
Self-reported diagnostic volume in 2005						Reference
< 250	3,227	5.7	85.7	5.4	3.3	2.37 (1.24–4.54)
250–499	4,471	3.8	92.3	2.3	1.6	1.76 (0.93–3.33)
500–999	3,990	3.8	92.8	1.7	1.7	1.83 (0.84–3.98)
≥ 1,000	2,972	4.4	90.7	4.1	0.8	
Portion of workload spent interpreting screening mammograms (%)						Reference
0–10	5,432	4.9	91.3	2.1	1.7	0.97 (0.61–1.54)
11–25	7,375	3.3	91.7	3.5	1.5	0.69 (0.39–1.21)
> 25	2,470	5.5	88.1	3.8	2.5	
Portion of workload spent interpreting diagnostic mammograms (%)						Reference
1–10	10,437	4.6	90.1	3.0	2.2	0.80 (0.32–1.98)
11–20	1,358	5.4	89.5	4.6	0.6	3.52 (1.18–10.49)
21–30	1,553	1.0	98.1	0.5	0.5	0.48 (0.13–1.82)
31–40	1,045	4.7	87.7	6.6	1.1	1.96 (1.06–3.61)
> 40	884	3.1	95.0	1.0	0.9	

^aThe number of examinations may not add to the total because of missing data.

^bOdds ratios (ORs) were adjusted for multiple examinations per radiologist, Breast Cancer Surveillance Consortium registry, age, density, and time since previous mammography. Bold ORs are statistically significant compared with the reference group.

^cAge and years of interpretation allowed to vary with year of examination.