Patient Compliance and Diagnostic Yield of 18-Month Unilateral Follow-Up in Surveillance of Probably Benign Mammographic Lesions

OBJECTIVE. The purpose of this study is to determine the patient compliance with and diagnostic yield of 18-month unilateral mammography in surveillance of probably benign (BI-RADS category 3) lesions.

MATERIALS AND METHODS. This retrospective study identified lesions prospectively classified BI-RADS 3 in asymptomatic women from January 1, 2004, to December 31, 2008. Surveillance protocol for BI-RADS 3 lesions included 6-month (unilateral), 12-month (bilateral), 18-month (unilateral), and 24-month (bilateral) imaging, with subsequent annual screening. Demographics, surveillance data, BI-RADS upgrades and downgrades, and biopsy results were abstracted from the longitudinal medical record.

RESULTS. One thousand one hundred eighty-eight lesions in 1077 patients (mean age, 51.5 years; age range, 26–89 years) had BI-RADS 3 assessment, representing 1.07% of all screening examinations. The compliance rates for follow-up at 6, 12, 18, and 24 months were 83.3%, 75.9%, 54.8%, and 53.9%, respectively. Sixty lesions were upgraded to BI-RADS 4 or 5 during surveillance. Biopsy revealed 15 cancers (cancer yield of 1.47%) from 1017 lesions with either 24-month imaging stability or tissue diagnosis available. Five, six, one, and three cancers were detected at 6, 12, 18, and 24 months, respectively. Cancers were all stage 0 or 1 except for one stage 2A cancer. Seven hundred forty-four of 1188 (62.6%) BI-RADS 3 lesions were downgraded before completing 2-year surveillance.

CONCLUSION. Most (11/15 [73%]) breast cancers initially assessed as BI-RADS 3 are diagnosed at up to 12 months’ surveillance. Eighteen-month unilateral mammography performed as BI-RADS 3 surveillance contributes minimally to cancer detection and has poor patient compliance.

Probably benign” is a clinically useful assessment category and is defined in the American College of Radiology’s BI-RADS lexicon as category 3 [1]. Lesions that are not typically benign but considered to have a very low likelihood of malignancy (< 2%) can be assessed as BI-RADS 3 and generally followed by imaging on a shorter interval—typically 6 months, rather than 1 year. Therefore, immediate tissue diagnosis can be avoided, and potential procedure-related complications and health care costs can be decreased.

Although short-interval follow-up for BI-RADS 3 assessment is a widely used practice among breast imaging specialists, there is a lack of consensus on the optimal follow-up protocol for surveillance, a problem addressed previously [2]. Published articles reporting outcome results of BI-RADS 3 lesions [3–5] predominantly adopted the protocol first advocated by Sickles [5], which includes a repeat unilateral diagnostic mammogram at 6 months and a bilateral diagnostic mammogram at 12 and 24 months; the 36-month examination may be optionally diagnostic or screening. Thus, although these are commonly referred to as “short-interval follow-up,” the only true short-interval examination during the surveillance period is the first 6-month unilateral follow-up examination. In the United States, where annual screening is recommended for women 40 years old and older [6], the other subsequent diagnostic examinations will coincide with the annual mammographic screening examinations. The protocol illustrated in the BI-RADS Atlas also follows this scheme and includes 6-, 12-, 24-, and optional 36-month follow-up examinations [1].

An alternative proposed protocol, also commonly used, includes follow-up diag-
Materials and Methods

This retrospective study was approved by our institutional review board. A search of our mammography database was performed to identify women who received a new, prospective assessment of a probably benign (BI-RADS 3) lesion after screening mammography during a 5-year period (January 1, 2004, to December 31, 2008). Our mammography database included our main teaching hospital, a cancer specialty center, and a large ambulatory outpatient imaging center. The examinations from the cancer center, encompassing high-risk patients in a surveillance program and patients currently or previously treated for breast cancer, were excluded to obtain a normal-risk population. During these 5 years, a total of 110,776 screening mammograms were performed, 69,033 at our teaching hospital and 41,743 at our ambulatory center. Of 1255 lesions that received a BI-RADS 3 assessment, 67 lesions were excluded for the following reasons: BI-RADS 3 assessment during diagnostic evaluation (clinical symptoms, history of breast cancer, or high risk) rather than screening examination (n = 53); immediate biopsy performed despite BI-RADS 3 assessment (n = 10); or referral from an outside institution with BI-RADS 3 assessment (n = 4). Therefore, the final study cohort included 1188 asymptomatic lesions (111 patients with bilateral BI-RADS 3 assessment). Nine hundred fifty-two (80%) BI-RADS 3 lesions were from the ambulatory center, and 236 (20%) lesions were from the teaching hospital. Thus, the overall rate of BI-RADS 3 assignment from screening was 1.07% (i.e., 1188/110,776) — 2.28% (952/41,743) at the ambulatory center and 0.34% (236/69,033) at the teaching hospital.

Mammograms were performed with digital technique on Senographe DS and Senographe Essential equipment (GE Healthcare). During the study time period, mammograms at the teaching hospital were interpreted by dedicated breast imagers, and mammograms at the ambulatory center were interpreted by general radiologists.

In our practice, diagnostic evaluation is always performed before assigning screening patients to BI-RADS 3. BI-RADS 3 assessments were assigned according to previously published criteria [1, 2, 4, 5] and included noncalcified circumscribed masses with round, oval, or gently lobulated contour; clustered round or punctate calcifications; focal asymmetries or one-view asymmetries; coarse calcifications suggestive of evolving dystrophic calcifications; multiple similar masses; or multiple similar groups of round calcifications. The usual follow-up imaging protocol for BI-RADS 3 lesions at our institution included 6-month (unilateral), 12-month (bilateral), 18-month (unilateral), and 24-month (bilateral) diagnostic imaging. When both breasts were assessed as BI-RADS 3, follow-up included bilateral diagnostic imaging every 6 months for up to 2 years. Because patients sometimes missed scheduled follow-up studies or rescheduled to a later date, follow-up studies were considered to be 6-, 12-, 18-, or 24-month examinations on the basis of timing relative to the initial BI-RADS 3 assessment, regardless of the number of the previously completed follow-up examinations. For instance, an examination from 18 to 23 months after initial BI-RADS 3 assessment was considered an 18-month follow-up; additionally, a patient who had examinations at 8, 19, and 27 months after the initial BI-RADS 3 assessment would have completed 6-, 18-, and 24-month follow-up by our definition. The compliance rates were calculated as the percentages of the total of 1188 lesions. Outcome data were assessed from the subcohorts who had documented 2-year imaging stability or tissue diagnosis during the 24-month surveillance period.

Review of the electronic medical record was performed to obtain the following information: patient age; lesion characteristics; date of initial BI-RADS 3 assessment; BI-RADS assessments for follow-up examinations, including BI-RADS upgrades and downgrades; duration of follow-up; and results of any biopsies, including staging information for any malignancies.

Results

Of the 1188 asymptomatic screening patients with 1188 lesions who received a BI-RADS 3 assessment, age ranged from 26 to 89 years (mean, 51.5 years). For 300 of the 1188 (25.3%) lesions, no prior comparison studies were available at the time of BI-RADS 3 assessment. Lesion type included 650 calcifications (54.7%), 261 masses (22.0%), and 302 focal asymmetries or one-view asymmetries (25.4%). Thirty-two lesions had two or more lesion types. Multiple findings leading to BI-RADS 3 assessment were counted as one lesion for the purpose of this study. Eleven (0.9%) lesions were assigned BI-RADS 3 on the basis of other findings, such as postsurgical appearance or diffuse increased bilateral density.

The BI-RADS assessments for each follow-up examination and patient compliance at each time point during 24 months are summarized in Figure 1. Of the 1188 total lesions, 990 (83.3%) returned at 6-month follow-up; 241 lesions were downgraded to BI-RADS 1 or 2, and 29 lesions were upgraded to BI-RADS 4 or 5, with biopsy yielding five cancers. Of the remaining 918 lesions still considered BI-RADS 3, 697 (75.9%) returned at 12-month follow-up; of these, 395 lesions were downgraded, and 14 lesions were upgraded, with biopsy yielding five cancers. Notably, one additional cancer diagnosed at 12 months (Fig. 2) was a BI-RADS 3 lesion that had been downgraded at the 6-month surveillance examination. Of the remaining 509 lesions still
considered BI-RADS 3, 279 (54.8%) returned at 18-month follow-up; 108 lesions were downgraded, and six lesions were upgraded, with biopsy yielding one cancer. Finally, of the remaining 395 lesions still considered BI-RADS 3, 213 (53.9%) returned at 24-month follow-up; 171 lesions were downgraded, and six lesions were upgraded, with biopsy yielding two cancers. Notably, a third cancer diagnosed at 24 months (Fig. 3) had been downgraded at 12-month follow-up and was upgraded to BI-RADS 4 at the 24-month screening study, with biopsy yielding cancer. There were three additional lesions that were previously downgraded and then upgraded to BI-RADS 4 on screening at 24 months with benign biopsies. The number of cancers diagnosed at each time point is shown in Figure 4.

Outcome data, defined as tissue diagnosis or at least 24 months of imaging stability, was known in 1017 lesions. Of the original 1188 lesions, 957 (including lesions downgraded to BI-RADS 1 or 2 before 24 months) were confirmed stable or resolved at 24 or more months and can be considered benign. Sixty lesions had tissue diagnosis before 24 months. One hundred seventy-one lesions were lost to follow-up before completing 24 months of imaging.

Overall, in the 2-year follow-up period, 60 of 1017 (5.9%) lesions with at least 24 months of imaging stability or a tissue diagnosis were upgraded, and 15 of these 1017 (1.47%) lesions were proved malignant. The positive predictive value of BI-RADS 3 lesions that were upgraded and biopsied was 25.0% (15/60). Reported imaging findings, follow-up surveillance data, and pathologic results of the cancers initially characterized as BI-RADS 3 lesions are detailed in Table 1. All 15 cancers were either stage 0 or 1, except for one stage 2A cancer. There were also eight incidentally detected cancers diagnosed during the 2-year surveillance period that were unrelated to the original BI-RADS 3 lesions—five within the breast undergoing surveillance and three in the contralateral screened breast.

The downgrade rate of lesions was 24.3% (241/990) at 6 months, 56.7% (395/697) at 12 months, and 38.7% (108/279) at 18 months. Of the total of 1188 lesions in the entire study group, 744 (62.6%) were downgraded to BI-RADS 1 or 2 before completing 2 years of surveillance. Only two cancers were ultimately detected from the downgraded 744 (0.27%) lesions. Typical reasons for downgrade included reassessment as benign by a different radiologist; follow-up of a cyst or lymph node; probably benign calcifications evolving to coarser, typically benign morphology (e.g., evolving fat necrosis or fibroadenoma); reactive lymph nodes; and one-view asymmetries dismissed as summation shadow on follow-up imaging.

Discussion

Recommending an 18-month follow-up mammogram as part of the short-interval follow-up protocol for BI-RADS 3 assessment is a common alternative to the surveillance protocol outlined in the BI-RADS lexicon [1]. Since a short-interval follow-up is generally considered to be a 6-month follow-up [13], some radiologists may extrapolate this protocol to include follow-up exams every 6 months for 2 or 3 years. This study is the first to our knowledge to investigate the diagnostic value of the 18-month follow-up in the management of probably benign mammographically detected nonpalpable breast lesions.

The 1.07% overall rate of BI-RADS 3 assessment from screening mammography in our study is somewhat lower than the BI-RADS 3 utilization reported by some other studies. A recent study using data from the Digital Mammographic Imaging Screen-
TABLE 1: Characteristics of Cancers Initially Classified as BI-RADS 3 Lesions

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age (y)</th>
<th>Site</th>
<th>Reported Imaging Finding(s)</th>
<th>Follow-Upa (mo)</th>
<th>Pathology</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>77</td>
<td>Amb</td>
<td>Scattered calcifications, some of which look vascular</td>
<td>6</td>
<td>DCIS, high grade</td>
<td>0(TisNXMX)</td>
</tr>
<tr>
<td>2</td>
<td>65</td>
<td>Hosp</td>
<td>New 4-mm circumscribed mass not seen on ultrasound</td>
<td>6, 12, 18</td>
<td>IDC, well differentiated</td>
<td>1A(T1bN0MX, 0.6 cm)</td>
</tr>
<tr>
<td>3</td>
<td>73</td>
<td>Amb</td>
<td>New 10-mm mass not seen on ultrasound</td>
<td>6</td>
<td>IDC, poorly differentiated</td>
<td>1A(T1cN0MX, 1.1 cm)</td>
</tr>
<tr>
<td>4</td>
<td>58</td>
<td>Amb</td>
<td>Punctate calcifications</td>
<td>6, 12</td>
<td>IDC, moderately differentiated</td>
<td>1A(T1bN0MX, 0.6 cm)</td>
</tr>
<tr>
<td>5</td>
<td>67</td>
<td>Amb</td>
<td>Fat-containing oval mass</td>
<td>6</td>
<td>IDC, moderately differentiated</td>
<td>1A(T1cN0MX, 1.2 cm)</td>
</tr>
<tr>
<td>6</td>
<td>53</td>
<td>Hosp</td>
<td>15-mm oval mass, probably fat necrosis</td>
<td>6, 12</td>
<td>IDC/ILC, moderately differentiated</td>
<td>1A(T1cN0MX, 1.4 cm)</td>
</tr>
<tr>
<td>7</td>
<td>57</td>
<td>Amb</td>
<td>Coarse calcifications</td>
<td>6</td>
<td>DCIS, high grade with microinvasion</td>
<td>1A(T1mN0MX)</td>
</tr>
<tr>
<td>8</td>
<td>59</td>
<td>Amb</td>
<td>Punctate calcifications</td>
<td>6</td>
<td>DCIS, high grade</td>
<td>0(TisNXMX)</td>
</tr>
<tr>
<td>9</td>
<td>59</td>
<td>Amb</td>
<td>Asymmetry not seen on ultrasound</td>
<td>6, 12</td>
<td>IDC, moderately differentiated</td>
<td>1A(T1aNXMX, 0.2 cm)</td>
</tr>
<tr>
<td>10</td>
<td>79</td>
<td>Amb</td>
<td>Focal asymmetry not seen on ultrasound</td>
<td>—, 12, —, 24</td>
<td>IDC, moderately differentiated</td>
<td>2A(T1cN1aMX, 1.4 cm)</td>
</tr>
<tr>
<td>11</td>
<td>61</td>
<td>Amb</td>
<td>Punctate calcifications</td>
<td>6, 12</td>
<td>IDC, moderately differentiated</td>
<td>1A(T1bN0MX, 0.9 cm)</td>
</tr>
<tr>
<td>12</td>
<td>42</td>
<td>Amb</td>
<td>Scattered coarse, punctate, and round calcifications</td>
<td>6, 12</td>
<td>IDC, poorly differentiated</td>
<td>1A(T1bN0MX, 1.0 cm)</td>
</tr>
<tr>
<td>13</td>
<td>40</td>
<td>Hosp</td>
<td>Grouped calcifications, probably vascular</td>
<td>6, 12, 18, 24</td>
<td>IDC, moderately differentiated</td>
<td>1A(T1aNXMX, 0.1 cm)</td>
</tr>
<tr>
<td>14</td>
<td>56</td>
<td>Amb</td>
<td>Scattered heterogeneous, punctate, and round calcifications</td>
<td>6, 12, —, 24</td>
<td>DCIS, high grade</td>
<td>0(TisNXMX)</td>
</tr>
<tr>
<td>15</td>
<td>53</td>
<td>Amb</td>
<td>4-mm pliable focal asymmetry not seen on ultrasound</td>
<td>—, 12</td>
<td>DCIS, low grade</td>
<td>0(TisNXMX)</td>
</tr>
</tbody>
</table>

Note — Some of the reported imaging findings would not qualify as probably benign according to the BI-RADS lexicon [1]. Hosp = hospital; Amb = ambulatory center, DCIS = ductal carcinoma in situ, IDC = invasive ductal carcinoma, IDC/ILC = invasive carcinoma with ductal and lobular features.

*Biopsies were performed promptly after the last time point shown. Dash indicates missed follow-up.

18-Month Follow-Up of Probably Benign Mammographic Lesions

Aging Trial pooled from 33 institutions during 2001–2003 reported a BI-RADS 3 rate of 2.34% [14]. Another study, using data from the Breast Cancer Surveillance Consortium during 1996–1999, reported BI-RADS 3 rates of 5.2% from first-round screening and 1.7% from subsequent screening examinations [15]. An older study, using data from the National Breast and Cervical Cancer Early Detection Program, reported an overall BI-RADS 3 rate of 7.7% during 1991–1996, which declined to 6.0% in 1996–1999 [16]. That study also reported a wide range of site-specific rates, from 1.1% to 12.2%. A similarly wide range of BI-RADS 3 assessments, from 1.2% to 9.8% among 40 centers, was reported using data from the Women’s Health Initiative [17], with an average of 5%. In our institution, our protocol is to assign a BI-RADS 3 assessment only after a full diagnostic workup, including additional mammographic views and ultrasound, as emphasized by many studies [1–4, 13]; also, many of our patients had prior imaging available. These two factors are known to decrease the BI-RADS 3 assessment rate [15, 16], and these factors may account in part for our lower rate of BI-RADS 3 assessment. Additionally, targeted breast ultrasound is used liberally in our practice and by most breast imaging centers currently, whereas it was less frequently used in the time period of older published studies. Targeted breast ultrasound can change a BI-RADS 3 assessment of a mammographic mass to BI-RADS 2 by showing a lymph node or cyst.

The difference in BI-RADS 3 assessments between our academic (0.34%) and community (2.28%) practices—even though they serve a similar normal-risk community—suggests that specialty training may also affect BI-RADS 3 utilization. Sickles et al. [18] found that, compared with general radiologists, specialist radiologists made more true-positive and fewer false-positive interpretations in screening mammography. In diagnostic mammography, specialists referred more patients for biopsy and found more malignancies than did general radiologists; thus, specialists had higher cancer detection rates in both screening and diagnostic mammography than did general radiologists. Another study evaluating the accuracy of short-interval follow-up (BI-RADS 3) mammograms found that radiologists spending more than 10 hours per week on breast imaging had higher sensitivity in detecting cancer during follow-up imaging than did radiologists spending less time on breast imaging [19]. Avoiding inappropriate BI-RADS 3 assessment can be challenging both for dedicated breast imagers at an academic institution [20] and for community-based radiologists.
[21]. In an unblinded retrospective review of BI-RADS 3 lesions ultimately proved malignant, Rosen et al. [20] found that none strictly met BI-RADS 3 criteria and most had already showed change at the time of BI-RADS 3 assignment. Similarly, some of the malignancies in Table 1 would not qualify for a probably benign lesion on the basis of the prospective imaging description, such as a new mammographic mass without an ultrasound correlate. Many factors undoubtedly contribute to BI-RADS 3 utilization, and a full analysis is beyond the scope of this study.

Similar to the results of Sickles [5], most cancers (73%) in our study initially classified as BI-RADS 3 after diagnostic evaluation were detected at the 6- or 12-month diagnostic study. Only one of 15 (6.7%) cancers in our study was detected at 18-month follow-up. The diagnostic yield at 18 months remains the lowest in the surveillance period even after taking into consideration the lower patient compliance at this time point. Our results suggest that the 18-month follow-up has a minimal role in detecting cancer. The potential benefit and cost of the 18-month follow-up can be also summarized as follows: detection of one early stage cancer 6 months earlier than the normally recommended 24-month annual follow-up [6] would be at the cost of 508 ultimately benign lesions requiring reassessment at 18 months.

Patient compliance with recommended imaging surveillance intervals was only moderate in our study. Although the compliance rate was initially high at 6-month (83.3%) and 12-month (75.9%) intervals, the return rate dropped to 54.8% and 53.0% at 18- and 24-month intervals, respectively. Poor patient compliance with BI-RADS 3 surveillance has also been reported by others [3, 14]. For the 18-month follow-up, if only the lesions that were also assessed at both 6- and 12-month follow-up are counted as truly compliant with the follow-up protocol, then the compliance rate drops even further, to 42.4%. For the 24-month follow-up time point, the similarly low compliance rate is even more problematic, because noncompliant patients on BI-RADS 3 surveillance at the 24-month time point are also failing to undergo contralateral annual screening. Our data prompt the concern that frequent surveillance every 6 months for 2 years (rather than a 6-, 12-, and 24-month protocol, which requires only one additional scheduled 6-month examination) may be negatively affecting the compliance of women who need routine screening of the other breast.

Short-term follow-up mammographic surveillance will occasionally identify additional incidental breast malignancies. In our study, eight incidental cancers unrelated to the original BI-RADS 3 lesion were detected during the surveillance period. Although more frequent imaging offers the potential for increased cancer detection, cost-effectiveness, radiation exposure, and potential effect on future compliance with imaging need to be considered.

The amount of downgrading of lesions to benign categories that occurred during the study period is interesting. In this study, 62.6% of the BI-RADS 3 lesions were downgraded to BI-RADS 1 or BI-RADS 2 at 6-, 12-, or 18-month follow-up, typically by a
different radiologist than the one who gave the initial BI-RADS 3 assessment. The downgrading does not appear to be inappropriate, given that only two cancers were detected among the downgraded lesions. The BI-RADS lexicon [1] states that more-experienced readers may downgrade the BI-RADS 3 assessment at follow-up examinations. Potential causes for the downgrades include interobserver variability in interpretation of the same finding and the finding’s evolving into a more typically benign appearance.

It has been stressed that BI-RADS 3 assessment is mainly reserved for interpreting baseline mammograms or those that do not have available prior studies [22]. In actual practice, BI-RADS 3 is used at times even when prior studies are available to the interpreting radiologist. When the prior studies provide a suboptimal comparison, including poor quality or different technique (i.e., analog to digital), resulting in probable but uncertain imaging stability, BI-RADS 3 may also be appropriate. In clinical practice, BI-RADS 3 may also occasionally be used in situations in which prior mammograms confirm mammographic change but the change is thought most likely to be benign; examples include suspected posttraumatic fat necrosis, calcifications suspected to be evolving dystrophic calcifications (particularly in a background of other clearly dystrophic calcifications), or findings thought to be more visible owing to technical reasons such as better compression, better technique, or fatty involution. In a breast imaging center with consistently high-quality imaging and a minority of patients presenting as a baseline or transitioning from other imaging centers, very low BI-RADS 3 utilization would be anticipated.

Our study has some limitations. This study is not part of a multicenter study, and it is possible that other breast imaging centers might have a different experience—especially in terms of compliance at 18 months, downgrade rate of lesions in surveillance, or both. Because this is a retrospective study, it is not possible to analyze the reasons for downgraded lesions during our study period. Only follow-up examinations at our institutions were available; any patient who went elsewhere for imaging would be lost to follow-up, and we cannot distinguish these patients from noncompliant patients. Our hospital and ambulatory care center were not completely separate independent imaging centers but rather were complementary; although patients tended to stay within one site for imaging follow-up, it was possible for patients to attend either center for imaging. Last, inappropriate BI-RADS 3 assessment for some of the lesions, as evident in Table 1, could adversely affect the BI-RADS 3 utilization rate and timely cancer diagnosis.

In summary, our data suggest that an 18-month follow-up unilateral mammogram performed as part of a BI-RADS 3 follow-up protocol contributes minimally to cancer detection and is hindered by poor patient compliance. Therefore, we recommend a 6-month (unilateral), 12-month (bilateral), and 24-month (bilateral) follow-up diagnostic mammography protocol for lesions assessed as BI-RADS 3 after a complete diagnostic evaluation. This protocol, as Sickles [2] pointed out, respects the annual screening routine of the patient with only one additional 6-month unilateral examination required, and it reduces the time and expense involved with imaging every 6 months for a 24-month surveillance period. Patient compliance with BI-RADS 3 surveillance remains a problem, which may be improved with this change in follow-up protocol. Breast imagers have the opportunity to improve patient compliance with short-interval follow-up imaging at 6, 12, and 24 months by ensuring adequate patient education regarding the importance of unilateral close surveillance and continued contralateral screening to maximize early breast cancer detection.

Acknowledgment

We are grateful to Nora McCarthy for assistance with data collection and data entry.

References

2. Sickles EA. Probably benign breast lesions: when should follow-up be recommended and is the optimal follow-up protocol? Radiology 1999; 213:11–14
12. Rubin E. Commentary on Dr. Sickles’ viewpoint. Radiology 1999; 213:21