Causes of Physician Delay in the Diagnosis of Breast Cancer

William H. Goodson III, MD; Dan H. Moore II, PhD

Background: Understanding sources of physician delay in diagnosis of breast cancer will assist efforts to expedite diagnosis.

Objective: To test whether increased reliance on screening mammography has affected causes of physician delay in diagnosis of breast cancer.

Design: Survey of delays in a case series.

Setting: Practice specializing in breast diseases in a region with high use of screening mammography.

Patients: Four hundred thirty-five consecutive patients treated for 454 breast cancers of any stage.

Intervention: Customary patient care.

Main Outcome Measures: Whether delay was related to how cancer was identified, patient age, individual cancer characteristics (such as tumor type), mammography reports, or physician expertise.

Results: Twenty-one women (5%) were inappropriately reassured that a malignant lump was benign without biopsy, 14 women (3%) had a misread mammogram, 4 women (1%) had a misread pathologic finding, and 5 women (1%) had cancer missed by a poorly performed fine-needle aspiration biopsy. Delay was associated with a benign mammography report (relative risk, 10.8; 95% confidence interval, 5.1-22.8), a woman finding her own mass (relative risk, 3.3; 95% confidence interval, 1.8-6.2), and current hormone replacement therapy (relative risk, 3.1; 95% confidence interval, 1.2-8.5).

Conclusions: The leading cause of physician delay in diagnosis of breast cancer continues to be inappropriate reassurance that a mass is benign without biopsy. Reducing delay in diagnosis will require less willingness to rely on clinical examination to decide that a mass is benign, less reliance on benign mammography reports to decide not to biopsy a mass, and a requirement that fine-needle aspiration biopsy be done by persons with demonstrated competence for the procedure.

Arch Intern Med. 2002;162:1343-1348
PATIENTS AND METHODS

From January 1, 1992, through December 31, 1999, we prepared short clinical abstracts from the medical records of 435 consecutive patients with 454 breast cancers of any stage referred from multiple sources to one surgeon (W.H.G.). For this analysis, bilateral cancer was counted as 2 cancers. Multiple foci of cancer in 1 breast was counted as 1 cancer. Information from these abstracts was put into a database (FileMaker Pro; Claris Corporation, Santa Clara, Calif) for analysis with standard statistical programs (Excel; Microsoft Corporation, Redmond, Wash; Data Desk; Data Description Inc, Ithaca, NY; and Stata, Stata Corp, College Station, Tex).

A detailed history was obtained from each patient at the initial visit. Chief complaint was defined as the first observation (by the patient, her physician, a radiologist, or other) that made the patient aware of the abnormality that was eventually diagnosed as cancer, without regard to how or by whom diagnostic tissue was obtained. We use diagnosis strictly to denote confirmation of cancer by histological or cytological examination. The history of whether the patient had been reassured that a mass was benign without biopsy was based on the patient’s understanding of what she had been told. Other information was corroborated by review of medical records, mammograms, histological slides, and other sources. We defined delay as physician action that completed an episode of care without diagnosing cancer of which there was a sign. Delay was tabulated if (1) a physician acknowledged a breast abnormality on physical examination but reassured the patient, without biopsy, that the mass was benign, and biopsy of that abnormality found cancer (offering a biopsy was not considered a mitigating action if biopsy had not been encouraged); (2) a fine-needle aspiration biopsy (FNA) obtained no malignant cells from a palpable or mammographic lesion, and cancer was found with a subsequent FNA by a physician with formal training in the procedure; (3) palpation-guided FNA obtained no malignant cells from a mammographic lesion, the patient was reassured, and subsequent image-guided biopsy found cancer; (4) a written pathology report indicated benign tissue and another pathologist diagnosed cancer on the same slides; or (5) a mammography report did not recommend biopsy or short-interval follow-up, review of the original x-ray films found suspicious calcifications or change, and biopsy confirmed cancer at that site. Length of delay was counted from physician action until diagnostic steps were resumed.

Physician delay was not tabulated if the responsible physician recommended a mammogram or a second physical examination after a short interval, if the physician recommended another diagnostic procedure, or if the original mammograms were unavailable.

As factors associated with delay in diagnosis, we examined chief complaint at the time of diagnosis, age, menopausal status, current hormone replacement therapy (HRT) in women after menopause, tumor type and grade, and interpretation of mammogram and FNA (if performed).

### RESULTS

Four hundred thirty-five women and 1 man underwent surgery for 454 cancers during the study (median age, 54 years; range, 24-88 years). Nineteen women had a second cancer in the contralateral breast. Data are analyzed as 454 separate cancers occurring in 454 separate breasts. No patient had delay in the diagnosis of both breasts. As shown in Table 1, the most common chief complaint for women younger than 50 years was the patient finding a mass, whereas for women 50 years and older, the most common chief complaint was a suspicious mammographic finding.

Delay occurred for 42 (9%) of 454 cancers. Nineteen (11%) of 175 cancers diagnosed before referral had delay, compared with 23 (8%) of 279 cancers diagnosed after referral ($\chi^2=0.87; P=.22$). Nineteen (45%) of 42 cancers with delay had diagnosis before referral, compared with 156 (38%) of 412 cancers without delay. The median delay was 9 months (range, 1-36 months), with a delay of 3 or more months for 30 patients and 6 or more months for 24 patients.

Twenty-one women (5%) were inappropriately reassured by their physician (15 primary care physicians and 6 consultant surgeons) that a lump was benign without biopsy. Fourteen (3%) had mammogram interpretation that failed to note a new mass (8 patients) or suspicious calcifications (6 patients). One mammography report noted a benign cyst but overlooked 5 cm of microlcalcifications. (The 2 patients with misread mammograms and inappropriate reassurance are counted in both categories; Table 2). Four patients had misread pathology reports from a surgical biopsy. Three had “benign” pathology reports, but review of the original histological slides revealed ductal carcinoma in situ. One of these women died of invasive cancer, identified by subsequent biopsy of residual calcifications in the same cluster. One had a diagnosis of lobular carcinoma in situ, but review of the original slides found ductal carcinoma in situ that required reexcision to obtain negative margins. Five women had poorly performed FNA by a physician without formal training in FNA technique, 3 by sur-
geons and 2 by other physicians. Four had no malignant cells obtained by FNA of a palpable mass that was subsequently diagnosed by FNA performed by a physician with specific training in FNA. Three of these women had delays of 4 to 6 months; 1 had workup resumed within a month when her overdue, routine mammogram showed suspicious microcalcifications. One patient with suspicious mammographic findings had benign results on a palpation-guided FNA of an area her surgeon assumed was the same as the area of suspicion on the mammogram, but subsequent image-guided biopsy found cancer, ie, her physician did an FNA in the wrong area.

None of the women with misread mammograms, misread pathologic findings, or poorly performed FNA were aware of delay before referral. Seventeen of 21 women with inappropriate reassurance were diagnosed when they sought further medical care (9 were diagnosed before referral; Table 2).

**UNIVARIATE ANALYSIS**

When women found their own lump, physician delay was 3 times more likely than when cancer was found by mammography, physician examination, or another sign (relative risk, 3.3; 95% confidence interval, 1.8-6.2; Table 3). Also among the postmenopausal group, 8 instances of inappropriate reassurance occurred in women taking HRT (χ²=9.3, P=.002). For women in the postmenopausal group who found their own lump, 8 of 26 taking HRT received inappropriate reassurance, compared with none of 34 not taking HRT (χ²=12.1; P<.001).

---

**Table 2. Chief Complaints at the Time of Diagnosis After Various Types of Initial Delays**

<table>
<thead>
<tr>
<th>Initial Cause of Delay</th>
<th>Inappropriate Reassurance</th>
<th>Misread Mammogram</th>
<th>Incorrect Reading of Pathology Report</th>
<th>Poorly Performed Fine-needle Aspiration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief Complaint*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient found mass</td>
<td>16</td>
<td>6</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Physician found mass</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Suspicious mammogram</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1†</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>12</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

*Leading to the detection of cancer at the time of diagnosis. Chief complaints for all patients are listed in Table 1.
†All women who received inappropriate reassurance had a palpable abnormality at the time of delay.
‡Random biopsy because of strong family history.

---

**Table 3. Univariate Analysis of Factors Possibly Relating to Delay**

<table>
<thead>
<tr>
<th>Factor</th>
<th>No. of Cancers</th>
<th>No. (%) With Delay</th>
<th>χ² Value</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief complaint</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient found mass</td>
<td>161</td>
<td>27 (16.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physician found mass</td>
<td>52</td>
<td>5 (9.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspicious mammogram</td>
<td>217</td>
<td>9 (4.1)</td>
<td>18.3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Other</td>
<td>24</td>
<td>1 (4.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal</td>
<td>454</td>
<td>42 (9.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>174</td>
<td>21 (12.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥50</td>
<td>280</td>
<td>21 (7.5)</td>
<td>2.6</td>
<td>.11</td>
</tr>
<tr>
<td>Subtotal</td>
<td>454</td>
<td>42 (9.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ductal</td>
<td>251</td>
<td>29 (11.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lobular</td>
<td>48</td>
<td>6 (12.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ductal and lobular</td>
<td>26</td>
<td>3 (11.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCIS</td>
<td>93</td>
<td>4 (4.3)</td>
<td>4.4</td>
<td>.22†</td>
</tr>
<tr>
<td>Subtotal</td>
<td>418</td>
<td>42 (10.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Written mammogram report</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign</td>
<td>121</td>
<td>32 (26.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspicious</td>
<td>312</td>
<td>8 (2.6)</td>
<td>59.3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Subtotal</td>
<td>433</td>
<td>40 (9.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor differentiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well</td>
<td>123</td>
<td>19 (15.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderately</td>
<td>115</td>
<td>12 (10.4)</td>
<td>2.7</td>
<td>.06‡</td>
</tr>
<tr>
<td>Poorly</td>
<td>83</td>
<td>7 (8.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal</td>
<td>321</td>
<td>36 (11.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRT in postmenopausal women</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No current HRT</td>
<td>137</td>
<td>5 (3.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current HRT</td>
<td>122</td>
<td>14 (11.5)</td>
<td>5.8</td>
<td>.02</td>
</tr>
<tr>
<td>Subtotal</td>
<td>259</td>
<td>19 (7.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*DCIS indicates ductal carcinoma in situ; HRT, hormone replacement therapy.
†If calculated as pooled invasive tumors vs DCIS, then χ² = 4.4; P = .04.
‡Does not include 12 DCIS with microinvasion, 3 mixed DCIS and lobular carcinoma in situ (LCIS), 16 LCIS, 2 inflammatory, 2 chest wall recurrences, and 1 malignant cystosarcoma.
§P value based on Cochran test for linear trend in proportions.
Delay was less likely for ductal carcinoma in situ than for invasive cancers. However, 66 (71%) of 93 ductal carcinomas in situ were identified by mammography vs 129 (40%) of 325 invasive cancers, so this difference may reflect expedited follow-up of abnormal mammographic findings rather than a factor predisposing to less delay.

There was a trend for more delays with low-grade tumors ($P = .06$). Delay was not associated with lobular vs ductal histological findings or age, nor was age associated with a specific cause of delay.

Physician delay was not associated with outcomes of larger tumor size, more positive nodes, or higher likelihood of having a mastectomy (Table 4). However, 4 of 24 women with more than 6 months’ delay filed malpractice claims against the responsible physician. These women were ages 40, 44, 47, and 55 at the time of diagnosis.

### MULTIVARIATE AND HIERARCHICAL ANALYSIS

In multivariate logistic analysis, a woman finding her own mass and taking HRT were statistically significant variables ($P = .004$ and $P = .049$, respectively). In a hierarchical model, benign mammographic findings and a woman finding her own mass were significant factors within the group of women taking HRT ($P < .001$ and $P = .03$, respectively). Among women who found their own mass, a benign mammographic finding was a significant factor ($P < .001$), and current use of HRT was marginally significant ($P = .06$).

### COMMENT

Despite widespread screening mammography, inappropriate reassurance that a palpable mass is benign, without biopsy, remains the leading cause of physician-delayed diagnosis of breast cancer, affecting 5% of patients in our series. Combining all causes, 9% of patients experienced delay.

Thirty-five percent (161/454) of cancers were first identified by patients themselves, but a disproportionate 64% (27/42) of the delays occurred when the patient first identified their own mass. This is consistent with analyses of physician delay$^{2,5,10}$ and malpractice lawsuits$^{16}$ which found that physician reassurance that a self-discovered mass was benign was a major cause of delay.

The known higher false-positive rates for breast self-examination may explain the association between delay and self-discovery. In randomized trials of breast self-examination, women who did breast self-examination found, in addition to cancer, many extra benign masses that required evaluation.$^{17,18}$ Physicians seem to intuit lower cancer risk for self-discovered masses and seek to reassure patients using CBE or mammography, without performing biopsy.

There are 2 possible responses when a patient identifies a breast mass. One approach is to interpret what the patient has felt and to reassure the patient without biopsy if the physician believes that the mass is benign. We do not know how often such reassurance is given in practice, but 21 of our patients experienced delayed diagnosis because of such misinterpretation of a mass by their physician. Seventeen of these women had mammograms with no signs of cancer even in retrospect, ie, they had false-negative mammograms. The potential for a negative mammogram to contribute to misinterpretation of CBE is an old finding.$^{1,3,6,16}$ Because of a trend to substitute mammography for CBE,$^{15,16}$ inappropriate reliance on mammography may become a greater problem.

The alternative is to encourage diagnostic tissue sampling from all palpable breast masses.$^{1}$ Fine-needle aspiration is the least invasive way to obtain tissue, but 1% of our women experienced delay because FNA was attempted by an untrained physician. Similarly, Ljung et al$^{19}$ found that tissue sampling by physicians without formal training in FNA technique missed 25% of cancers, while physicians with formal training missed only 2% of cancers. The National Cancer Institute$^{20}$ Consensus Panel recommended that FNA be done only by persons with special training. Often this person is not a surgeon, but examination by a surgeon is desirable before needle biopsy, which might cause a hematoma. If no FNA-trained physician is available, surgical or core biopsies are alternatives.

Clinical breast examination is an inexact science, and there may be legitimate questions of whether a mass exists. Such judgments are difficult. Haagensen,$^{21}$ in the era before widespread mammography, described the “humiliating experience” of being shown a mass by a patient after his own detailed examination. A second examination after a short interval allows a fresh perspective to decide whether a mass exists and would have provided a second chance for prompt diagnosis for the 21 women who were inappropriately reassured. This is an accepted plan,$^{21-23}$ especially for women before menopause, whose breasts change cyclically.

Breasts also have varying areas of nodularity and density, and there may be limited, defined situations in which it is appropriate to reassure patients without biopsy. Such...
changes must be symmetrical in both breasts or be anticipated in the context of life events, such as prior biopsy. For example, breast tissue is typically more dense in the upper outer quadrant. If both breasts are symmetrically more dense in this area, this is likely the patient’s normal condition. Similarly, nodularity may increase after weight loss or when hormone effects subside, eg, after pregnancy. These changes are not focal, as are most cancers, and they do not explain a single, self-discovered mass.

Like Tarttter et al, who also studied various chief complaints, we did not find a relation of age to delay. However, malpractice litigation was filed only by 4 younger women, which agrees with the observation that malpractice plaintiffs are younger than the average woman with breast cancer. Prior reports linking delay to younger age included mainly women with self-discovered or palpable masses, rather than a spectrum of chief complaints, which likely explains the difference.

Like Richards et al, we found a trend for less delay among women with high-grade tumors. As Coates suggested, this may be because high-grade tumors are clinically more obvious, perhaps demonstrating greater tissue reaction to the cancer, and receive expedited care.

More frequent delay with HRT, particularly inappropriate reassurance, is a new observation. Hormone replacement therapy decreases sensitivity of mammography, probably because increased tissue density hides early signs of cancer. However, sensitivity of CBE is not the issue when inappropriate reassurance is given for a mass that has already been discovered. We suspect that knowledge of the effects of HRT on mammographic density is somehow extrapolated into a supposed explanation of why a mass is present in a given patient.

All diagnostic tests for breast cancer have known rates of misinterpretation. In our series, 217 cancers were first identified by mammography, but 14 (7%) had delay because of misinterpretation of mammograms. In other studies, reading of mammograms by a single radiologist missed 5% to 10% of cancers. Similarly, pathologists misread 1% of our cases, and others report misinterpretation of pathologic findings for 0.7% to 4% of breast cases. Review by a second radiologist or pathologist reduces misinterpretation of these tests.

Misinterpretation is also a known characteristic of CBE, as observed in previous surveys. In our series, 21 women had misinterpretation of CBE, 6 times by consultant surgeons. In a direct assessment of ability to classify masses, Hicks et al found that 18 of 197 masses classified as benign on CBE were malignant.

Unlike mammography and pathologic examination, however, second interpretation of a CBE does not greatly improve accuracy. When Chamberlain et al tried to increase accuracy, they found that, when 2 physicians examined 215 women with a family history of breast cancer, 4 of 6 cancers were in women whom both physicians agreed not to refer for biopsy. Similarly, the Yorkshire Breast Cancer Group found only 59% agreement between clinicians on whether to categorize masses as malignant or benign.

For precisely the reason that interpretation of CBE is an inexact science, once a mass is identified, there must be the equivalent of a second reading, but neither mammography nor another CBE is sufficient. Only tissue sampling is adequate. If a mass is followed up without biopsy, patients should be informed that interpretation of CBE has a known rate of misinterpretation even with an experienced examiner. Unless tissue sampling is encouraged, we view simply mentioning a biopsy as nothing more than self-protecting behavior by the physician.

Because this was a single practice, we cannot comment on the true frequency of delay in our community. However, because the proportion with delay was similar among women diagnosed before and after referral, and because the proportion of patients referred after diagnosis was similar in the delay and the nondelay groups, we know of no reason to suspect biased referral of a particular type of delay, and identification of factors associated with delay should not be affected. Furthermore, our data set has allowed us to investigate novel factors in delay, such as different chief complaints, the role of FNA, and current use of HRT.

These delays began when 42 different physicians used tests—interpretation of CBE, single interpretation of mammograms and pathologic findings, and FNA by untrained (but otherwise experienced) persons—that have known rates of inaccuracy. The performance characteristics of these tests (discussed herein) predict delay will occur at random in practices whenever these tests are widespread in the medical community. Therefore, we believe that most delayed diagnoses reflect not failures by individuals but, rather, the overall inadequacy of practices that are accepted, despite the fact that they fail even in experienced hands. Reeducating a few physicians will not eliminate delayed diagnosis. Rather, reducing delay will require collective changes in all clinicians’ perception of whether to interpret CBE, the reliability of mammography, whether to demand that tests like FNA be performed well or simply performed, and possibly how to interpret the secondary effects of HRT. To reduce physician delay, we suggest that physicians provide care with “FERVR”:

1. Focus CBE on the presence or absence of a mass, without interpretation of what it might be.
2. Expect tissue sampling for all palpable breast masses.
3. Recommend tissue sampling of all palpable masses even if the mammogram shows no signs of cancer.
4. Verify that physicians doing FNAs on their patients have specific training in the procedure.
5. Reevaluate patients after a short interval if a tentative decision is made that an area identified by a patient falls within the narrowly defined rules of predictable variations of breast structure.

Accepted for publication October 22, 2001.

This study was presented in part at the Second European Breast Cancer Conference, Brussels, Belgium, September 29, 2000.

We thank Brian Mayall, MD, and Stefanie Jeffery, MD, for constructive criticism during the preparation of the manuscript.

Corresponding author and reprints: William H. Goodson III, MD, Department of Surgery, California Pacific Medical Research Institute, 2100 Webster, Suite 401, San Francisco, CA 94115 (e-mail: whg3@cooper.cpmc.org).
REFERENCES