Prescreening Mammography by Technologists: A Preliminary Assessment

OBJECTIVE. We assessed the ability of technologists to accurately classify screening mammograms as either showing negative findings or requiring follow-up.

SUBJECTS AND METHODS. In a prospective study, 33 technologists at a central facility and five satellite breast imaging facilities recorded whether mammograms obtained during 3019 examinations showed negative findings or findings that indicated that additional procedures were required. The technologists were not specifically trained for the experiment. The technologists’ interpretations were compared with radiologists’ interpretations.

RESULTS. Technologists and radiologists agreed in 82% of the cases (77% negative findings and 5% requiring follow-up). Of the 175 cases recommended for follow-up by only the radiologists, 17 were ultimately biopsied and two were found to be malignant.

CONCLUSION. Even without undergoing additional training, technologists can perform at reasonable levels of accuracy in classifying screening mammograms. The possibility of using technologists to group cases after the technologists have undergone training is an interesting concept that should be explored further.

The American Cancer Society recommends that all women older than 40 years undergo periodic screening mammography for the purpose of detecting breast cancer early; as a result, millions of procedures are performed each year [1–3]. The steady increase since 1987 in the number of screening mammograms obtained each year in the United States is evidence that compliance is improving [4]. However, the number of qualified radiologists with mammography training is not following this trend [5]. Some studies performed in the United States and abroad indicate that specially trained nonphysicians can efficiently and effectively contribute to the diagnostic process of interpreting mammograms.

One study, performed by Hillman et al. [6], found that specifically trained physician assistants could interpret mammograms with a high degree of accuracy. Most other studies in this area have trained X-ray technologists to evaluate mammograms. Accuracy levels in these studies vary, but these findings, which were obtained under restricted laboratory environments, are encouraging [7–9].

Although somewhat outdated perhaps, the concept of using nonphysicians to assist in the diagnostic process remains an interesting one because of the rapid changes in the clinical practice of mammography screening. Mammography technologists have frequent and direct access to the acquired images, and many technologists have a large amount of experience reviewing mammograms with both negative and abnormal findings for quality assurance purposes. Therefore, with additional training, technologists could accurately group screening mammograms into two basic categories: cases that require further evaluation and cases that have either negative or clearly benign findings and the patient needs to return only at the next periodic screening date.

The purpose of this study was to assess in a prospective manner the baseline ability (i.e., before training) of mammography technologists to classify screening mammograms into these two groups.

Subjects and Methods

Methodology

The study was performed at Magee-Womens Hospital, a large urban hospital affiliated with the University of Pittsburgh Medical Center Health System, and at five of its community satellite breast imaging facilities. In 2000, more than 29,300 screening mammograms were obtained at these centers...
mammography examinations were performed at these facilities with a cancer detection rate of 2.7 cases per 1000 examinations.

All mammography technologists from the hospital and the satellite facilities were asked to participate in this institutional review board–approved study to evaluate their ability to classify mammograms as either requiring follow-up or not requiring follow-up. The 33 technologists who participated did not receive formal training for the purpose of this study. At the time of the study, the number of years that the technologists had worked specifically in mammography ranged from 2 to 26. The technologists had a median of 13 years of experience, and approximately 60% are certified by the American Cancer Society as instructors of breast self-examination.

The technologists were asked to take a minute or two to evaluate the screening mammograms they acquired when they performed general quality assurance. Because the technologists review available previously obtained mammograms as part of their routine, the technologists had the opportunity to look at previously obtained mammograms, when available, during the assessment. Viewing conditions were not controlled during the study (e.g., room lighting and background masking). After reviewing a screening mammogram, the technologists completed a short form that included the patient’s and technologist’s identification numbers, the date of the examination, where the examination was performed, and the category into which they classified the mammograms. The two possible categories were as follows: the patient requires additional workup, such as additional views, sonography, or biopsy; or the patient may return in a year without further workup. Technologists were not asked to identify the specific reason for the recommendation for additional workup or the specific location of a suspected abnormality, when applicable. They were not instructed to err on the false-negative side to increase sensitivity. Data collection was performed over a period of approximately 6 weeks, yielding 3019 examination ratings.

All forms were collected by a department supervisor and submitted to the research staff. Data were entered into a database in which the technologists were identified only by a number, which was listed with the examination date, a patient identifier, and a binary classification code. Data about all cases were input into a second database; in addition, the interpretation made by one of nine radiologists was ascertained from the reports (0, recommending further interpretation made by one of nine radiologists was ascertained from the reports (0, recommending further interpretation made by one of nine radiologists was ascertained from the reports (0, recommending further diagnosis). We also examined the reasons for recall provided by the radiologists in their reports so that we could determine which features were interpreted as requiring follow-up by the radiologists but not by the technologists.

Cases were initially divided into five categories: suspicious microcalcifications, suspicious nodule or mass, asymmetry, palpable mass not visualized on mammograms or the acquisition protocol, cases of suspected ruptured implant, regions of scar tissue, and cases with multiple reasons for recall. Because all breast examinations at our facilities (screening and diagnostic) include a clinical examination and because a sonogram and additional views (or both) of a palpable abnormality—regardless of whether it is depicted on the images—are often obtained, we listed these cases in the initial distribution as reported by the radiologists; however, these cases were later excluded from the analysis because follow-up was recommended on the basis of information provided from another source, not imaging. Technical recalls, as determined by the radiologists, were also excluded from the analysis.

To investigate the possibility that technologists, particularly before receiving special training, are more likely to characterize a mammogram as showing suspicious findings if a benign abnormality is present, we analyzed the data for the subset of cases: those that had been classified by the radiologists as a category 2 lesion, according to the Breast Imaging Reporting and Data System (BI-RADS) [11].

### Results

Table 1 shows the total number of cases with agreement and disagreement between the classifications made by radiologists and technologists after exclusion of all cases with palpable findings (n = 21) and cases that were recalled for technical reasons (n = 13). Of 2985 cases, radiologists and technologists agreed about whether mammograms showed negative findings (2284 cases or 77% of the total population) or findings that required follow-up (157 cases or 5% of the total population). Disagreement occurred in 18% of the cases: in 12% of the cases, only technologists recommended follow-up, and in 6%, only the radiologists did so. In seven of the 2985 examinations, malignancy was verified.

Table 2 shows, as an example, the distribution of cases with agreement and disagreement between radiologists and technologists for 11 technologists who made 100 or more assessments each. Of these 11 technologists, three had a disagreement rate of less than $5\%$ of the cases recommended only by the radiologists for a follow-up (technologists 1, 7, and 9). Furthermore, the highest miss rate for this group was 10% for technologist 10 (with an average of the number of cases per 10,000 interpretations.

### Table 1

<table>
<thead>
<tr>
<th>Radiologists’ Interpretations</th>
<th>Technologists’ Interpretations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative Findings</td>
<td>Total</td>
</tr>
<tr>
<td>Negative findings</td>
<td>2284 (77)</td>
</tr>
<tr>
<td>Follow-up required</td>
<td>175 (6)</td>
</tr>
<tr>
<td>Total</td>
<td>2459 (83)</td>
</tr>
<tr>
<td>Follow-up required</td>
<td>369 (12)</td>
</tr>
<tr>
<td>Total</td>
<td>526 (17)</td>
</tr>
<tr>
<td>Total</td>
<td>2985 (100)</td>
</tr>
</tbody>
</table>

Note.—The excluded cases included 13 cases in which a technical problem occurred and 21 palpable masses. The numbers in parentheses represent the percentage of the total cases in each category.
Prescreening Mammography by Technologists

of 6%), if we assume that the radiologists’ interpretations are correct (i.e., a gold standard). Similar results were observed for the group as a whole. However, although participation was completely voluntary, this group of technologists clearly showed, partially through self-selection, an active interest in participating. We report the level of agreement between the technologists’ interpretations and the radiologists’ interpretations.

Table 3 shows the distribution of cases with agreement and disagreement between radiologists and technologists for radiologists who made more than 100 assessments each in this group of cases. Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases). Of note are the average recall rate of 13% (for this group of cases).

The agreement between the recommendations of radiologists and those of technologists was high, particularly when considering that the technologists did not receive any special training to perform the required classification task. This high level of agreement results, in part, from the fact that most of the findings in cases in this study were actually negative. As important perhaps is the fact that the recommendation rate for follow-up by technologists was only 50% higher than that of the radiologists. Technologists identified most of the cases that were later proven to be malignant as requiring additional follow-up, and more than 80% of the cases that they classified were in agreement with the radiologists. In addition, the radiologists carefully reviewed findings from previously performed examinations if available, whereas the review of previously performed examinations by the technologists for quality assurance was not as rigorous. The low cancer rate detected in this group should

<table>
<thead>
<tr>
<th>Radiologist No.</th>
<th>Total No. of Cases Evaluated</th>
<th>Recalled by</th>
<th>Not Recalled by Either</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>915</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>566</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>460</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>248</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>232</td>
<td>19</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>223</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>173</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>Average</td>
<td>402</td>
<td>13</td>
<td>6</td>
</tr>
</tbody>
</table>

TABLE 3 Percentage of Cases Recommended for Follow-Up by Radiologists Who Evaluated More Than 100 Cases

would generally be expected. Finally, because of the large volume of cases in our practice and the variability in the technologists’ abilities to make independent decisions, all decisions about acquiring additional images are made by physicians; technologists are limited to acquiring only four views per examination unless a technical fault clearly exists in the study. Hence, in general, our radiologists are more concerned about false-negative findings for subtle cases.

Of the 832 cases rated by the radiologists as a BI-RADS category 2 lesion (i.e., a benign finding) in at least one breast, technologists recommended follow-up for 183 (22%). In contrast, they recommended only 17% of all other cases for follow-up. Of the cases recommended for follow-up by only the radiologists (n = 175 patients), additional views (one bilaterally) were obtained in 96 patients, and of this group, nine underwent biopsy. Sonography was recommended for 43 patients, and one underwent biopsy. Both additional views and sonography were recommended for 36 patients. Seven of these 36 patients ultimately underwent biopsy. Seventeen biopsies were performed at our institution in this group of patients: two were found to be malignant and 15 were benign. One of the malignant lesions was depicted on mammography as a subtle cluster of microcalcifications, and the other was suspected because the radiologist detected architectural distortion on mammography.

Discussion

The agreement between the recommendations of radiologists and those of technologists was high, particularly when considering that the technologists did not receive any special training to perform the required classification task. This high level of agreement results, in part, from the fact that most of the findings in cases in this study were actually negative. As important perhaps is the fact that the recommendation rate for follow-up by technologists was only 50% higher than that of the radiologists. Technologists identified most of the cases that were later proven to be malignant as requiring additional follow-up, and more than 80% of the cases that they classified were in agreement with the radiologists. In addition, the radiologists carefully reviewed findings from previously performed examinations if available, whereas the review of previously performed examinations by the technologists for quality assurance was not as rigorous. The low cancer rate detected in this group should

TABLE 2 Percentage of Cases Recommended for Follow-Up by Technologists Who Evaluated More Than 100 Cases

<table>
<thead>
<tr>
<th>Technologist No.</th>
<th>Total No. of Cases Evaluated</th>
<th>Recalled by</th>
<th>Not Recalled by Either</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>316</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>267</td>
<td>19</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>226</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>187</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>181</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>150</td>
<td>24</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>142</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>138</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>136</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>126</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>112</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>Average</td>
<td>180</td>
<td>16</td>
<td>10</td>
</tr>
</tbody>
</table>
not be alarming because the average age of the patients was relatively low and many live in residential areas. We recognize that some of the patients included in our screening schedule, such as a patient with an implant, may undergo only diagnostic mammography in other radiology practices. However, the number of these cases in our study was small and would not affect the main findings.

Despite these encouraging results, using technologists as physician extenders in the diagnostic process would clearly require that they be specifically trained for this purpose. Both their detection sensitivity and specificity must increase above the baseline levels reported here. The wide distribution of reporting in terms of cases per technologist stems from several factors, including work schedule and individual assignments to different areas (e.g., diagnostic, screening, biopsy). Because we did not force technologists to report every screening case and during the period in question we performed approximately 5000 screening procedures, intentional selection of cases may have occurred. Namely, some technologists may have reported only what they perceived to be “easy” or “clear” cases. However, because of the general level of compliance in recordings (~60% of all cases were reported), we suspect that most participants did not intentionally select cases. Our preliminary results indicate that accurate determination of benign findings and recognition of potentially important asymmetry are only some of the areas in which training could improve the technologists’ performance. Even if technologists can be trained to detect suspected abnormalities, characterizing the abnormalities as benign or malignant with a high degree of accuracy requires a substantial amount of training. This project addressed only the former step in detection.

We recognize that the approach we investigated in this preliminary study has little impact on current practices. Our study was designed as a prospective real-time experiment that did not substantially affect technologists’ workflow and did not affect the radiologists’ workflow. Only a few studies of this type have, to our knowledge, been reported in the literature. By its nature, this design yielded limitations that must be recognized. Most important, cases were not specifically selected; hence, the number of cases with positive findings was small, and the number of errors, as shown by the estimated false-negative rate, was large. Lastly, any attempt to formally include technologists in the diagnostic process clearly should address a number of difficult issues including, but not limited to, the handling of the litigious nature of mammography.

---

**Sumkin et al.**

in which training could improve the technologists’ performance. Even if technologists can be trained to detect suspected abnormalities, characterizing the abnormalities as benign or malignant with a high degree of accuracy requires a substantial amount of training. This project addressed only the former step in detection.

We recognize that the approach we investigated in this preliminary study has little impact on current practices. Our study was designed as a prospective real-time experiment that did not substantially affect technologists’ workflow and did not affect the radiologists’ workflow. Only a few studies of this type have, to our knowledge, been reported in the literature. By its nature, this design yielded limitations that must be recognized. Most important, cases were not specifically selected; hence, the number of cases with positive findings was small, and the number of errors, as shown by the estimated false-negative rate, was large. Lastly, any attempt to formally include technologists in the diagnostic process clearly should address a number of difficult issues including, but not limited to, the handling of the litigious nature of mammography.

---

**References**


This article has been cited by:


3. Jolanda Meijers. 18 Mammae 791-830. [Crossref]


8. Louise M. Henderson, Thad Benefield, Mary W. Marsh, Bruce F. Schroeder, Danielle D. Durham, Bonnie C. Yankaskas, J. Michael Bowling. 2015. The Influence of Mammographic Technologists on Radiologists’ Ability to Interpret Screening Mammograms in Community Practice. *Academic Radiology* **22**:3, 278-289. [Crossref]


22. WM Thompson, K Pollard. 2007. The current status of radiographers as screen readers in breast screening units: applications for Australia. *Radiographer* **54**:3, 16-20. [Crossref]


27. N.Reed Dunnick. 2004. ACR intersociety conference 2003: radiologist assistants and other radiologist extenders. *Journal of the American College of Radiology* **1**:6, 386-391. [Crossref]