Screening for Breast Cancer

Joann G. Elmore, MD, MPH
Katrina Armstrong, MD
Constance D. Lehman, MD, PhD
Suzanne W. Fletcher, MD, MSc

Breast cancer screening, especially with mammography, has been recommended for many decades, and the majority of women older than 40 years in the United States participate in screening activities. Meanwhile, new screening modalities have been introduced, and some of these have been increasingly incorporated into community practice. However, none of the new technologies has been evaluated for its effect on breast cancer mortality.

Community practice of screening may differ from the care provided within randomized clinical trials and is less often discussed in review articles. Reviews of breast cancer screening usually emphasize efficacy and results of randomized trials, particularly those involving screen-film mammography.

Efficacy of a screening tool is measured in randomized trials assessing effectiveness of breast cancer screening were reviewed, as well as meta-analyses, systematic reviews, studies of breast cancer screening in the community, and guidelines. Also, studies of newer screening modalities were assessed.

Data Sources and Study Selection

All major US medical organizations recommend screening mammography for women aged 40 years and older. Screening mammography reduces breast cancer mortality by about 20% to 35% in women aged 50 to 69 years and slightly less in women aged 40 to 49 years at 14 years of follow-up. Approximately 95% of women with abnormalities on screening mammograms do not have breast cancer with variability based on such factors as age of the woman and assessment category assigned by the radiologist. Studies comparing full-field digital mammography to screen film have not shown statistically significant differences in cancer detection while the impact on recall rates (percentage of screening mammograms considered to have positive results) was unclear. One study suggested that computer-aided detection increases cancer detection rates and recall rates while a second larger study did not find any significant differences. Screening clinical breast examination detects some cancers missed by mammography, but the sensitivity reported in the community is lower (28% to 36%) than in randomized trials (about 54%). Breast self-examination has not been shown to be effective in reducing breast cancer mortality, but it does increase the number of breast biopsies performed because of false-positives. Magnetic resonance imaging and ultrasound are being studied for screening women at high risk for breast cancer but are not recommended for screening the general population. Sensitivity of magnetic resonance imaging in high-risk women has been found to be much higher than that of mammography but specificity is generally lower. Effect of the magnetic resonance imaging on breast cancer mortality is not known. A balanced discussion of possible benefits and harms of screening should be undertaken with each woman.

Conclusions In the community, mammography remains the main screening tool while the effectiveness of clinical breast examination and self-examination are less. New screening modalities are unlikely to replace mammography in the near future for screening the general population.

Context Breast cancer screening in community practices may be different from that in randomized controlled trials. New screening modalities are becoming available.

Objectives To review breast cancer screening, especially in the community and to examine evidence about new screening modalities.

Data Sources and Study Selection English-language articles of randomized controlled trials assessing effectiveness of breast cancer screening were reviewed, as well as meta-analyses, systematic reviews, studies of breast cancer screening in the community, and guidelines. Also, studies of newer screening modalities were assessed.

Data Synthesis All major US medical organizations recommend screening mammography for women aged 40 years and older. Screening mammography reduces breast cancer mortality by about 20% to 35% in women aged 50 to 69 years and slightly less in women aged 40 to 49 years at 14 years of follow-up. Approximately 95% of women with abnormalities on screening mammograms do not have breast cancer with variability based on such factors as age of the woman and assessment category assigned by the radiologist. Studies comparing full-field digital mammography to screen film have not shown statistically significant differences in cancer detection while the impact on recall rates (percentage of screening mammograms considered to have positive results) was unclear. One study suggested that computer-aided detection increases cancer detection rates and recall rates while a second larger study did not find any significant differences. Screening clinical breast examination detects some cancers missed by mammography, but the sensitivity reported in the community is lower (28% to 36%) than in randomized trials (about 54%). Breast self-examination has not been shown to be effective in reducing breast cancer mortality, but it does increase the number of breast biopsies performed because of false-positives. Magnetic resonance imaging and ultrasound are being studied for screening women at high risk for breast cancer but are not recommended for screening the general population. Sensitivity of magnetic resonance imaging in high-risk women has been found to be much higher than that of mammography but specificity is generally lower. Effect of the magnetic resonance imaging on breast cancer mortality is not known. A balanced discussion of possible benefits and harms of screening should be undertaken with each woman.

Conclusions In the community, mammography remains the main screening tool while the effectiveness of clinical breast examination and self-examination are less. New screening modalities are unlikely to replace mammography in the near future for screening the general population.
those of randomized clinical trials. In addition, we reviewed what is known about newer screening modalities, specifically digital mammography, computer-aided detection programs for mammography, ultrasound, and magnetic resonance imaging (MRI).

METHODS

The evaluation of screening modalities, especially in the community setting, is challenging for methodological, clinical, and ethical reasons. Randomized clinical trials are considered the gold standard for evaluating a new screening test. The long-term breast cancer mortality rate of women randomized to receive a new screening test is compared with that of women randomized to receive standard care. However, such trials are difficult to conduct. They require tens of thousands of women who need to be followed up for more than 15 years. Furthermore, because mammography screening has been shown to be effective in some trials, it would likely be even more difficult to demonstrate any additional efficacy of new tests. Finally, as treatment for breast cancer has improved over time, the impact of screening on breast cancer mortality may be increasingly difficult to establish.

Because of these challenges, new screening tests are often first studied by establishing characteristics of the tests themselves, rather than by studying their effect on patient outcome such as breast cancer mortality. Important test characteristics include sensitivity, specificity, safety, cost, simplicity, and patient and clinician acceptability. We review what test characteristics have been studied and the findings for each new modality. We also indicate the study design and the endpoints studied for each screening test. Although it is important to use resources wisely when considering a screening test for a large segment of the population, cost-effectiveness analyses are not reviewed.

It is important to determine the characteristics of a screening test in a community setting if the test is to be used in that setting. However, test characteristics of new modalities are usually evaluated among women for whom the rate of breast cancer is higher than average, such as women at increased risk of breast cancer or women in a diagnostic setting with breast symptoms or known breast abnormalities. The reported sensitivity and specificity of a test in these high-risk women may be different from the sensitivity and specificity of the same test used in a general screening population. We therefore indicate if a test has been evaluated as a diagnostic or screening test and if as a screening test, whether it has been evaluated in women thought to be at increased risk or in the general population.

For this review, searches of MEDLINE, The Cochrane Library, the National Guideline Clearinghouse Web site, the US Preventive Services Task Force recommendations and reviews, and the International Agency for Research on Cancer Handbook of Cancer Prevention (IARC) were performed to identify English-language articles about breast cancer screening. Examples of search terms included mass screening and breast and then the specific modality (e.g., mammography, breast examination, digital mammography, computer assisted diagnosis or computer aided detection, magnetic resonance imaging, ultrasound or ultrasonography). The bibliographies of retrieved articles were also scanned to retrieve additional relevant articles. Details regarding search terms are available from the authors.

RESULTS

Screening Mammography

Eight reported randomized trials have studied mammography’s effectiveness in the United States, Sweden, Canada, and the United Kingdom. Concerns related to flaws of these randomized clinical trials have been raised. In-depth independent reviews of the criticisms of the trials have concluded that these flaws do not negate mammography’s efficacy in reducing breast cancer mortality, especially in women aged 50 to 69 years. Trials comparing mammography with or without clinical breast ex-
BREAST CANCER SCREENING

Table 1. Breast Imaging Reporting and Data System Assessment Categories Used in the United States for Mammography Examinations and Associated Likelihood Ratio for Breast Cancer Diagnosis*

<table>
<thead>
<tr>
<th>Assessment Category</th>
<th>Assessment</th>
<th>Definition</th>
<th>Likelihood Ratio for Breast Cancer Diagnosis†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Negative</td>
<td>Breasts appear normal</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>Benign finding</td>
<td>A negative mammogram result, but the interpreter wishes to describe a finding</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>Probably benign finding; short-interval follow-up suggested</td>
<td>Lesion with a high probability of being benign noted on mammogram</td>
<td>1.2</td>
</tr>
<tr>
<td>0</td>
<td>Need additional imaging evaluation</td>
<td>A lesion is noted for which additional imaging evaluation is needed; used almost always in a screening situation</td>
<td>7.0</td>
</tr>
<tr>
<td>4</td>
<td>Suspicious abnormality—biopsy should be considered</td>
<td>A lesion is noted for which the radiologist has sufficient concern to recommend a biopsy</td>
<td>125</td>
</tr>
<tr>
<td>5</td>
<td>Highly suggestive of malignancy; appropriate action should be taken</td>
<td>A lesion is noted that has a high probability of being cancer</td>
<td>2200</td>
</tr>
</tbody>
</table>

*Categories also include an assessment of 6, used when there is a known diagnosis of breast cancer before the mammogram.
†Likelihood ratios for risk of breast cancer diagnosis at first screening mammography.

*Radiography were highest in older women, especially those older than 80 years. Screening mammography may not be beneficial for women with significant comorbidity or with a life expectancy of less than 5 years, and may actually necessitate work-up that does not result in any benefit. However, if a woman is in reasonably good health and would be eligible for and interested in treatment, continued screening should be supported. Encouraging individualized decisions may be especially appropriate for women older than 70 years.

The majority of women with abnormalities noted on screening mammograms (~95%) do not have breast cancer with variability based on multiple factors including the radiologist’s assessment and the woman’s age. Because the risk of breast cancer increases with age, the likelihood of a woman with an abnormal mammogram result having cancer also increases with age. On the other hand, having a normal mammogram result does not rule out the possibility of having breast cancer, because false-negative mammography examination results do occur. In such cases, either the cancer is not visible on mammography examination or the radiologist fails to notice the lesion prospectively.

Results from 7 population-based community screening programs in the United States on 463,372 screening mammograms revealed an overall sensitivity of 75.0% and specificity of 92.3%. A sensitivity of 75% means that 25% of women (25 of 100 women) who were diagnosed with breast cancer had normal mammogram results between 12 and 24 months before their cancer diagnosis (eg, false-negative examination result). This sensitivity of 75% in the community was similar to that reported in the randomized trials (68% to 88%) but the specificity was lower than in most of the trials (range, 82%-93% for Canadian National Breast Screening Study 1 to 98.5% for Health Insurance Plan of New York).

Breast density and age are important predictors of accuracy. Adjusted sensitivity ranged from 63% in women with extremely dense breasts to 87% in women with almost entirely fatty breasts; adjusted sensitivity increased with age from 69% in women aged 40 through 49 years to 83% in women aged 80 through 89 years. Adjusted specificity increased from 89% in women with extremely dense breasts to 97% in women with almost entirely fatty breasts.

Guidelines for quality assurance have been issued by several bodies, such as the Commission of the European Communities and the US Mammography Quality Standards Act. The Breast Imaging Reporting and Data System, used in the United States to standardize reports, includes categories for assessment ranging from 0 to 5 (Table 1). The associated likelihood ratios for a breast cancer diagnosis for first screen are shown in the table. Use of the Breast Imaging Reporting and Data System has not eliminated the variability among radiologists that had been noted before.

Large differences have been noted between the recall rates (or percentage of screening mammograms considered as positive) of community-based mammography programs in the United States and those in other countries. The recall rate in the United States is twice the recall rate in the United Kingdom (eg, 12.5%-14.4% vs 7.6%), with no difference in cancer detection rate. Elmore et al noted comparable differences between North American screening programs and those in other locations, which persisted after adjusting for differences such as age of women screened, use of single vs double reading, and use of 1 vs 2 views of each breast for examinations. Other possible reasons for the regional variability noted include differences in the characteristics of the population screened (eg, presence of risk factors or symptoms) and features of the mammography examination (eg, equipment type and year, technician training). The experience of the physician interpreting the mammograms has also been raised as a possible reason for the variability. Recommendations vary regarding the minimum number of mammograms that the physician should interpret yearly, from 480 in the United States to 5000 in the United King-

©2005 American Medical Association. All rights reserved.
dom.51 Educational training has been shown to improve sensitivity with no change in specificity.63 Finally, features of the health care system, including malpractice concerns, financial incentives, quality control procedures, and auditing procedures, may be related to some of the variability.60

Double-reading of films, where 2 or more radiologists interpret each film, is offered in some US screening programs and in about half of the other countries that use mammography screening.4,58 However, double reading is performed in various ways (eg, 2 interpretations may be completely independent, 2 radiologists may perform interpretations together, or a third reader may serve as a tie-breaker). A systematic review of 10 cohort studies found that double reading increased the cancer detection rate by 3 to 11 per 10000 women screened; recall rates increased or decreased depending on the method of double reading used.64 The heterogeneity of the definitions of double reading and of screening practices precludes an accurate assessment of benefits and harms.

Full-Field Digital Mammography and Computer-Aided Detection Programs

Both screen-film mammography and full-field digital mammography use x-rays to obtain images. With screen-film mammography the image is captured on film; with full-field digital mammography the image is captured digitally. Digital images can then be printed on film for viewing, or the images can be interpreted directly on a computer monitor. The digital acquisition process improves logistics and work flow by allowing electronic transmission, storage, and retrieval. Radiologists viewing the image on a monitor can alter the contrast and brightness of the image and magnify specific areas without additional x-ray exposure. Digital imaging also allows easier use of computer-aided detection software. One disadvantage, however, is the increased cost associated with full-field digital mammography.4

Three community-based studies have compared full-field digital mammography to screen-film mammography (Table 2). Two studies found the sensitivity of full-field digital mammography (64% and 74%) to be less than that of screen-film mammography (79%, 90%), but these studies had a small number of women with breast cancer (42 and 31, respectively) and the display systems and experience of radiologists may have improved since these studies.65,66,70 A larger randomized study reported similar cancer detection rates (per all screened), with higher recall rates for full-field digital mammography.67 A trial now being conducted, which aimed to enroll 49400 women, will compare the diagnostic accuracy of full-field digital mammography from 4 different manufacturers with that of traditional screen-film mammography.71

Computer-aided detection programs, which recognize patterns in breast images associated with cancer (Figure 1), may potentially help radiologists improve their diagnostic accuracy, but presently data are limited (Table 2).68,69,72,73 Computer programs that can mark calcifications, masses, or other potential lesions on the mammogram may increase the number of cancers detected compared with unassisted interpretations. In a study of 12860 women,68 computer-aided detection increased radiologists’ overall screening recall rate from 6.5% to 7.7% while increasing the number of cancers detected from 41 without computer-aided detection to 49 with the technology. However, in the largest clinical series to date, which included 59139 mammograms interpreted with computer-aided detection and 56432 interpreted without, recall rates and cancer detection rates did not differ significantly.69 Computer-aided detection may prove helpful in reducing variability among radiologists of differing expertise. Such programs are used by a small but growing number of mammography facilities in the United States.69,74 Medicare and Medicaid allow for additional billing for computer-aided interpretations of mammograms. As technologies continue to improve, larger multisite studies will be needed for more definitive evidence.

Clinical Breast Examination

Although two thirds of US women older than 40 years receive regular screening clinical breast examinations,2 few data about the efficacy of clinical breast examinations alone are available from randomized clinical trials. Four randomized trials of mammography included the clinical breast examination in the screened group.19,21-23 One of these trials, the Canadian National Breast Screening Study 2 of women aged 50 through 59 years at entry, compared the results of an annual standardized 10- to 15-minute clinical breast examination and breast self-examination with the results of an annual standardized clinical examination and breast self-examination plus mammography (in other words, there was no typical control group that received no screening).21 The trial found that breast cancer mortality was similar in the 2 groups of women although yearly mammography in addition to physical examination and breast self-examination detected more small and lymph node–negative breast cancers than did screening with physical examination alone.21 Although the sensitivity of screening clinical breast examination was highest at 63% in the National Breast Screening Study 2,73 an overall estimate based on all randomized trials calculated sensitivity at 54% (95% confidence interval [CI], 48%-60%) and specificity at 94% (95% CI, 90%-97%).76

We suspect that few clinicians in the community setting perform examinations as carefully as those in the Canadian trial, so accuracy may be lower in the community setting. Results reported from community practices showed sensitivity ranging from 28% to 36%77-79 (Table 3). In one study, two fifths of physicians (34 out of 80) who performed a screening breast examination on manufactured breast models used no discernible systematic search pattern at all.80 Sensitivity of examinations improved by spending more time80,91 and by using a thorough, sys-
tematic technique. However, the number of false-positive examinations may increase with training.

In randomized controlled trials, noting an abnormality on a screening clinical breast examination in an asymptomatic average-risk woman increased the likelihood of breast cancer (likelihood ratio [LR], 10.6; 95% CI, 5.8-19.2). However, in community practice, an abnormal screening breast examination result was associated with an LR of 2.1, substantially lower than that of women in the same practice presenting with a breast abnormality (LR, 24). Noting a suspicious abnormality on a screening mammogram was associated with LRs ranging from 7 to 2200 (Table 1).

**Breast Self-examination**

Breast self-examination is appealing as a patient-centered, noninvasive procedure that allows women to become comfortable with their own bodies. However, the extent of current practice is thought to be low. About one third of US women regularly perform breast self-examination, and the estimated sensitivity is low (20% to 30%). Among respondents to the Women Physicians’ Health Study (N = 4501), only 21% reported performing monthly breast self-examination.

Training in breast self-examination, while associated with increased accuracy of detection of lumps in breast tissue, has been associated with increased rates of false-positive findings and thus diminished specificity. In addition, there is evidence casting doubt on the benefits. A large randomized controlled trial in Shanghai, China, of 26604 women working in textile factories provided half of the women with intensive initial instruction, including practice with breast models, as well as regular reminders and practice examinations under supervision once every 6 months for 5 years. This study found no positive effect of breast self-examination on breast can-

## Table 2. Studies of Full-Field Digital Mammography and Computer-Aided Detection Programs in Community Screening Settings

<table>
<thead>
<tr>
<th>Source/Location</th>
<th>Population/Design</th>
<th>Screen Film Digital</th>
<th>Recall Rate Screen Film Digital</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lewin et al.65, 66 United States</td>
<td>6736 Examinations for 4489 women Women received both examinations</td>
<td>33/42 (78.6) 27/42 (64.3)</td>
<td>1007/6736 (14.9) 799/6736 (11.9)</td>
<td>2 Institutions</td>
</tr>
<tr>
<td>Skaane et al.66 2003 Norway</td>
<td>3683 Women Women received both examinations</td>
<td>28/31 (90.3) 23/31 (74.2)</td>
<td>128/3683 (3.5) 168/3683 (4.6)</td>
<td>8 Radiologists Initial recall rate before consensus meeting and comparison with previous mammograms available was 7% for screen film and 12% for digital</td>
</tr>
<tr>
<td>Skaane and Skjennald,67 2004 Norway</td>
<td>25263 Women Randomized to 1 examination</td>
<td>...</td>
<td>3.0 (45-49 y) 3.7 (45-49 y)</td>
<td>8 Radiologists 4 Independent radiologists interpreted all FFDM images for first 4 mo Initial recall rate before consensus meeting and comparison with prior examinations available was 9.8%-10.2% for screen film and 13.7%-14.0% for digital</td>
</tr>
<tr>
<td>Freer and Ulisssey,68 2001 United States</td>
<td>12860 Women Each mammogram was initially interpreted without CAD, followed by reevaluation of areas marked by CAD</td>
<td>3.2/1000 3.8/1000</td>
<td>6.5 7.7</td>
<td>2 Radiologists</td>
</tr>
<tr>
<td>Gur et al.69 2004 United States</td>
<td>56,432 Examinations without CAD 59,159 Examinations with CAD Pre-post study design after introduction of CAD into a clinical practice setting</td>
<td>3.49/1000 3.55/1000</td>
<td>11.39 11.40</td>
<td>24 Radiologists No adjustment for patient characteristics</td>
</tr>
</tbody>
</table>

Abbreviations: CAD, Computer-aided Detection; FFDM, full-field digital mammography; ellipses indicate that data are not provided on cancer detection among the subgroup of women with cancer (sensitivity). Data presented on cancer detection rates among all women screened were similar for digital and screen film.
cer mortality after 10 years of follow-up but almost double the rate of biopsies due to false-positive findings (1.8% of women in the instruction group vs 1.0% of women in the control group). The study results should be interpreted with caution because approximately 40% of women in the trial were in their 30s. Also, it is possible that a 10-year follow-up was not long enough to see an effect on breast cancer mortality.

A meta-analysis of the effect of regular breast self-examination on breast cancer mortality or rates of advanced breast cancer (a marker of death) was performed on 20 observational studies and 3 clinical trials. Bias and confounding may affect the results of studies of women with breast cancer who reported practicing self-examination before diagnosis. No difference in death rate was noted in studies of women who detected their cancers during self-examination (pooled relative risk [RR], 0.90; 95% CI, 0.72-1.12), and no mortality differences were noted in trials of training (pooled RR, 1.01; 95% CI, 0.92-1.12).

Magnetic Resonance Imaging

Although screen-film mammography and full-field digital mammography are the only imaging tools explicitly approved or grandfathered in for breast cancer screening by the US Food and Drug Administration (FDA), other modalities are under study. Those approved by the FDA for diagnostic purposes (not screening) include MRI, ultrasound, scintimammography, thermography, and electrical impedance imaging. Although mammography uses x-ray and sonography uses sound waves to create images, MRI produces images from the combination of a strong magnetic field, radio waves, and computer processing (FIGURE 2).

Screening MRI may be helpful for women for whom mammography is not optimal, such as young women at substantially increased risk for breast cancer because of known BRCA1 or BRCA2 mutations. Available data are limited to studies of test characteristics in women at high risk (TABLE 4) and the impact on breast cancer mortality has not been determined. Both retrospective and prospective cohorts have been described. The small number of cases with breast cancer in these studies (the range among studies was 3 to 45 women) means that estimates of sensitivity were not precise. Nevertheless, every study reported higher sensitivity for breast MRI than for mammography, ultrasound, or both. The largest study re-

Table 3. Studies of Screening Clinical Breast Examination Performance in Clinical Trials and Community Settings in the United States

<table>
<thead>
<tr>
<th>Population (Reference)</th>
<th>No. of Women Screened</th>
<th>Participants and Setting</th>
<th>Years</th>
<th>% (95% Confidence Interval)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pooled analysis of clinical trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6419 With cancer (429/248 with no cancer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants in randomized clinical trial (Health Insurance Plan of NY), Canadian Breast Screening Study, and nonrandomized controlled trial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1963-1988</td>
<td>54.1 (48.3-59.8)</td>
<td>94.0 (90.2-96.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| US National Breast and Cervical Cancer Early Detection Program |
| 589,048 |
| Low-income women (age ≥ 40 y) enrolled in National Breast and Cervical Cancer Early Detection Program |
| 1995-1998 | 36.1 | 96.2 |

| US Health Plan in Pacific Northwest |
| 468 |
| Health plan enrollees (age ≥ 40 y) in western Washington attending breast cancer screening program |
| 1988-1994 | 35.3 | Not available |

| Columbia Medical Center, Department of Radiology |
| 11,130 (27,825 Screens) |
| Urban women attending single academic screening program |
| 1995-2000 | 27.6 | 99.4 |

*Figure 1. Computer-Aided Detection Markers of Breast Imaging Screening Mammography*

Examination with subtle focal asymmetric density on mammography examination of the right breast, mediolateral oblique view. The density is marked by the computer-aided detection program, with a star to call attention to a possible cancer. Breast biopsy confirmed infiltrating ductal carcinoma. The triangle notes calcification in the same breast.

*Table 3. Studies of Screening Clinical Breast Examination Performance in Clinical Trials and Community Settings in the United States*
ported on 1909 women at increased risk in the Netherlands, with 45 women diagnosed with cancer who had all screening examinations. The sensitivity of clinical breast examination, mammography, and MRI in this study was 17.9%, 40%, and 71%, respectively. The overall discriminating capacity of MRI was significantly better compared with mammography as assessed by receiver operating characteristic curves (area under the curve 0.827 for MRI vs 0.686 for mammography).

Specificity of MRI tends to be lower than that of mammography; however, data are not consistently presented and specificity is not always easy to calculate. In the study of 1909 women in the Netherlands, the specificity of clinical breast examination, mammography, and MRI was 98.1%, 95.0%, and 89.8%, respectively, and the authors noted that screening with MRI led to twice as many unneeded additional examinations (420 vs 207) and 3 times as many unneeded biopsies (24 vs 7) as did screening with mammography. Warner et al reported a substantial recall rate in the first round of MRI screening (26%), which decreased to 10% in the third round of MRI screening. Additional studies of MRI screening among high-risk women are under way.

Magnetic resonance imaging has not been studied in the general population as a screening tool, and the results from MRI screening of high-risk women may not apply to women at average risk. The high cost of MRI (approximately 10 times the cost of mammography) and its relatively low specificity (compared with mammography) probably prohibit its use in screening.

Table 4. Comparative Results of Nonrandomized Studies of Screening Mammography, Magnetic Resonance Imaging, and Ultrasound in Screening Women at Increased Risk for Breast Cancer

<table>
<thead>
<tr>
<th>Source</th>
<th>Site</th>
<th>Study Design</th>
<th>Total No. of Women</th>
<th>Age, y Mean (Range)</th>
<th>No. (%) Known Mutation Carriers*</th>
<th>Sensitivity</th>
<th>Cancer Yield From MRI Alone</th>
<th>Biopsies Performed as a Result of MRI</th>
<th>PPV of Biopsies Performed Based on MRI, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kuhl et al.</td>
<td>Germany</td>
<td>Prospective</td>
<td>192</td>
<td>39 (18-65)</td>
<td>3/9 (33)</td>
<td>9/9 (100)</td>
<td>3/9 (33)</td>
<td>6/192 (3)</td>
<td>14/192 (7.3)</td>
</tr>
<tr>
<td>Tilanus-</td>
<td>The Netherlands</td>
<td>Prospective</td>
<td>109</td>
<td>42 (22-68)</td>
<td>12 (11)</td>
<td>0</td>
<td>3/3 (100)</td>
<td>. . .</td>
<td>3/109 (2.8)</td>
</tr>
<tr>
<td>Linthorst et</td>
<td>The Netherlands</td>
<td>Prospective</td>
<td>105</td>
<td>46 (25-77)</td>
<td>. . .</td>
<td>1/8 (13)</td>
<td>8/8 (100)</td>
<td>1/8 (13)</td>
<td>7/105 (6.7)</td>
</tr>
<tr>
<td>Podo et al.</td>
<td>Italy</td>
<td>Prospective</td>
<td>367</td>
<td>50 (23-82)</td>
<td>19 (5)</td>
<td>NA</td>
<td></td>
<td>14/14 (100)</td>
<td>14/367 (3.8)</td>
</tr>
<tr>
<td>Morris et al.</td>
<td>United States</td>
<td>Retrospective</td>
<td>236</td>
<td>47 (26-65)</td>
<td>236 (100)</td>
<td>8/22 (36)</td>
<td>17/22 (77)</td>
<td>7/21 (33)%</td>
<td>7/236 (3)</td>
</tr>
<tr>
<td>Warner et al.</td>
<td>Canada</td>
<td>Prospective</td>
<td>1909</td>
<td>40 (19-72)</td>
<td>358 (19)</td>
<td>18/45 (40)</td>
<td>32/45 (71)</td>
<td>22/1909 (2.2)</td>
<td>56/1909 (2.9)</td>
</tr>
<tr>
<td>Kriege et al.</td>
<td>The Netherlands</td>
<td>Prospective</td>
<td>1909</td>
<td>40 (19-72)</td>
<td>358 (19)</td>
<td>18/45 (40)</td>
<td>32/45 (71)</td>
<td>22/1909 (2.2)</td>
<td>56/1909 (2.9)</td>
</tr>
</tbody>
</table>

Abbreviations: MRI, magnetic resonance imaging; NA, not applicable; PPV, positive predictive value. Ellipses indicate that data were not available.

*Data are shown for women with known positive results.
†Reported PPV for 105 women with valid 1-year of follow-up.
‡To detect 3 breast cancers in 109 women, the investigators performed 193 MRI examinations, 51 ultrasounds, 29 fine-needle aspiratory cytology, and 2 benign excision biopsies.
§One patient who had cancer detected only by MRI did not receive ultrasound.
¶All women in this study had a normal mammogram result for inclusion.
©2005 American Medical Association. All rights reserved.
Box. Breast Cancer–Related Information

Calculation of the risk of a breast cancer diagnosis and death at the level of individual women:
http://bcra.nci.nih.gov/brc/start.htm (Gail model)
http://astor.som.jhmi.edu/brcapro (Gail Model, Claus Model, and a model that predicts the probability of carrying a BRCA1 or BRCA2 mutation)
http://yourcancerrisk.harvard.edu/
Breast Self-Examination Tutorials
http://www.komen.org/bse
http://www.breastsellexam.ca
National Guidelines for Breast Cancer Screening
http://www.guidelines.gov
Randomized Clinical Trials of New Modalities in Breast Cancer Screening
http://www.clinicaltrials.gov
http://www.acr.in/current_protocols.html

routine use for screening general populations. Also, MRI is time-consuming, requires intravenous contrast administration, and may be problematic for claustrophobic patients.

Ultrasound

Ultrasound, frequently used as a targeted diagnostic examination focusing on a specific area of concern, may help distinguish between cyst and solid masses and also between benign and malignant masses. Breast ultrasound data are available from diagnostic populations, with screening studies limited to women with dense breasts on mammography or at increased risk for breast cancer. Although ultrasound may detect 3 to 4 additional breast cancers per 1000 women in these increased-risk populations, there are no data on the use of screening ultrasound in the general population. Breast ultrasound has limitations as a potential screening tool because it requires a well-trained skilled operator. Examination techniques are not standardized, interpretation criteria are variable, and breast ultrasound does not consistently detect microcalcifications. Preliminary data suggest a higher rate of false-positive examination results with ultrasound than with mammography alone. For example, the false-positive rate (based on solid lesion for ultrasound) ranged from 2.4% to 12.9% for ultrasound and 0.7% to 6% for mammography.

National Screening Guidelines

All groups recommend screening mammography for women aged 50 through 69 years. Within the United States, all recommend it for women in their 40s, but vary in the screening intervals recommended and encourage “informed decision making” with all women about the choice. Being older than 70 years should not preclude women from continuing to undergo screening; however, decisions regarding continued screening should include life expectancy and health status. International policies differ with respect to the target age group to be screened, the intervals between screening, the number of mammographic views taken per breast, and the screening modalities recommended. Clinical breast examination is recommended by some, but not all, groups. Most national groups no longer recommend breast self-examination, but some encourage women to become familiar with the contour of their own breasts. Other imaging modalities, such as MRI and ultrasound, are not recommended for screening the general population. (See http://www.Guidelines.gov.)

Benefits and Harms

The primary goal of breast cancer screening is to reduce subsequent breast cancer mortality through early detection. Theoretically this should translate into reduced morbidity from the disease. In addition, many women report feeling reassured by screening, especially after having a so-called normal screen result.

Possible harms include pain and discomfort, especially noted during compression of breast tissue during mammography. Compression of breast tissue reduces motion artifact and improves image quality. Reports of the level of discomfort, however, vary widely. Anxiety about screening is another concern. Breast cancer screening yields both false-positive and false-negative results. False-positive results have been associated with anxiety, additional costs, and morbidity. After 10 years of annual screening in the United States, it is estimated that 1 in 2 women will have at least 1 false-positive mammogram result, and 1 in 5 women will have at least 1 false-positive clinical breast examination result. False-negative mammography examinations occur in approximately 20% to 40% of women with breast cancer.

Overdiagnosis and overtreatment of clinically insignificant disease is possible, especially ductal carcinoma in situ noted by mammography. Theoretical concerns about radiation-induced breast cancer from exposure to repeated mammography have been raised, but the potential benefits are thought to outweigh the risks. For example, the benefit-to-harm ratio is estimated to be 48.5 lives saved per 1 life lost due to radiation exposure. A mortality paradox has been noted in women aged 40 through 49 years, whereby increased mortality is noted among women screened for the first 3 to 10 years after the initiation of screening. Tumor dissemination after needle biopsy has also been suggested although the clinical significance is unclear.

Observer variability among radiologists who interpret mammography examinations has been noted both in a test situation and in community practices. A decision to perform a breast biopsy may depend heavily on the radiologist’s interpretation; therefore, interpretive variability can directly affect patient management. The benefit-to-harm ratio of screening increases as women age because screening accuracy improves and prevalence of breast cancer increases. Younger women, however, have more potential years of life to be gained from screening.
Communicating With Patients

Effective communication of information on benefits and harms is challenging. Multiple studies document inaccurate or incomplete comprehension of risk information, cognitive biases that affect how patients process risk information, and poor communication skills on the part of physicians. Use of frequencies with specific reference groups ("23 out of 1000 women your age") instead of percentages ("0.23%") may facilitate the comparison of small risks. Presentation of both positive ("23 in 1000 women your age will develop breast cancer") and negative framing ("977 in 1000 women your age will not develop breast cancer") can reduce biases in decision making. Visual aids, such as bar graphs or pie charts, can increase the comprehension and saliency of information. Risk-prediction models can be used to calculate the probability of being diagnosed with breast cancer (Box). Effective physician-patient partnership, for which gaps in comprehension are frequently assessed and resolved may have the greatest impact on patients’ understanding of information.

CONCLUSION

Reviews of breast cancer screening usually concentrate on the results of randomized trials of mammography to reduce breast cancer mortality. We have emphasized data on effectiveness in the community setting among the general population of women, which can often be different from the ideal setting of a randomized trial or the setting of a study among high-risk women. We have also emphasized the challenge of evaluating new screening modalities. Newer screening tests such as MRI and ultrasound have been studied in women at increased risk of breast cancer (eg, carriers of BRCA1 or BRCA2 mutations). None of the newer tests has been evaluated for its effect on breast cancer mortality in the general population and no data support screening the

### Table 5. Summary of Breast Cancer Screening Modalities for General Average-Risk Populations

<table>
<thead>
<tr>
<th>Screening Modality</th>
<th>Test Characteristics Investigated in the General Population by Type of Study Design</th>
<th>Effect on Breast Cancer Studied</th>
<th>FDA Approval</th>
<th>Medicare Reimbursement, $*</th>
<th>HCPCS Codes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screen-film mammography</td>
<td>Yes Yes Yes Yes†</td>
<td>Yes</td>
<td>Yes†</td>
<td>90 76 092</td>
<td>G0101</td>
<td>Most extensively studied screening modality</td>
</tr>
<tr>
<td>Clinical breast examination</td>
<td>Yes Yes Yes NA</td>
<td>Yes</td>
<td>NA</td>
<td>39 G0202</td>
<td>NA</td>
<td>Higher sensitivity associated with longer duration of examinations Sensitivity noted in community setting appears lower than trials</td>
</tr>
<tr>
<td>Breast self-examination</td>
<td>Yes Yes Yes NA</td>
<td>Yes</td>
<td>NA</td>
<td>NA NA</td>
<td>NA</td>
<td>No positive effect on mortality noted and increased biopsies because of false-positive findings</td>
</tr>
<tr>
<td>Digital mammography</td>
<td>Yes Yes No Yes</td>
<td>No</td>
<td>Yes†</td>
<td>144 G0202</td>
<td>NA</td>
<td>Improves logistical practice, facilitates use of computer-aided detection programs Comparable cancer detection rate to screen-film mammography The impact on recall rates is unclear</td>
</tr>
<tr>
<td>Computer-aided detection</td>
<td>No Yes No Yes</td>
<td>No</td>
<td>Yes†</td>
<td>21 76 083</td>
<td>NA</td>
<td>Cost in addition to that of mammography May help radiologists improve accuracy but no data to date to definitively support</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>No No No No</td>
<td>No</td>
<td>No</td>
<td>74 76 645</td>
<td>NA</td>
<td>Requires skilled operator, examination technique not standardized, interpretation criteria variable, does not detect microcalcifications</td>
</tr>
<tr>
<td>Magnetic resonance imaging</td>
<td>No No No No</td>
<td>No</td>
<td>No</td>
<td>1108 76 094</td>
<td>NA</td>
<td>More sensitive and slightly less specific than mammography in women at increased risk No radiation Requires intravenous contrast, time-consuming, may not be feasible for some women, such as those with pacemaker, aneurysm clips, or claustrophobia</td>
</tr>
</tbody>
</table>


*Represents the 2004 Medicare reimbursement rate for Seattle, Wash. The cost for computer-aided detection is in addition to the cost for the mammography examination.†Screen-film mammography was grandfathered in by the FDA.
43. Ma L, Fishel E, Wright B, Hanna W, Allan S, Boyd NF. Case-control study of factors associated with fail-
BREAST CANCER SCREENING


Downloaded From: http://jama.jamanetwork.com/pdfaccess.ashx?url=/data/journals/jama/4967/ on 06/18/2017