Background: This systematic review is an update of evidence since the 2002 U.S. Preventive Services Task Force recommendation on breast cancer screening.

Purpose: To determine the effectiveness of mammography screening in decreasing breast cancer mortality among average-risk women aged 40 to 49 years and 70 years or older, the effectiveness of clinical breast examination and breast self-examination, and the harms of screening.

Data Sources: Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews (through the fourth quarter of 2008), MEDLINE (January 2001 to December 2008), reference lists, and Web of Science searches for published studies and Breast Cancer Surveillance Consortium for screening mammography data.

Study Selection: Randomized, controlled trials with breast cancer mortality outcomes for screening effectiveness, and studies of various designs and multiple data sources for harms.

Data Extraction: Relevant data were abstracted, and study quality was rated by using established criteria.

Data Synthesis: Mammography screening reduces breast cancer mortality by 15% for women aged 39 to 49 years (relative risk, 0.85 [95% credible interval, 0.75 to 0.96]; 8 trials). Data are lacking for women aged 70 years or older. Radiation exposure from mammography is low. Patient adverse experiences are common and transient and do not affect screening practices. Estimates of overdiagnosis vary from 1% to 10%. Younger women have more false-positive mammography results and additional imaging but fewer biopsies than older women. Trials of clinical breast examination are ongoing; trials for breast self-examination showed no reductions in mortality but increases in benign biopsy results.

Limitation: Studies of older women, digital mammography, and magnetic resonance imaging are lacking.

Conclusion: Mammography screening reduces breast cancer mortality for women aged 39 to 69 years; data are insufficient for older women. False-positive mammography results and additional imaging are common. No benefit has been shown for clinical breast examination or breast self-examination.

Primary Funding Source: Agency for Healthcare Research and Quality.
abnormal mammographic finding on screening or a concern regarding finding on physical examination, additional imaging and biopsy may be recommended. Additional imaging may consist of diagnostic mammography or mammography done with additional or special views, targeted breast ultrasonography, or breast MRI (13, 14). Additional imaging may help classify the lesion as a benign or suspicious finding to determine the need for biopsy. Biopsy techniques vary in the level of invasiveness and amount of tissue acquired, which affects yield and patient experience.

We focus on new studies and evidence gaps that were unresolved at the time of the 2002 USPSTF recommendation. These include the effectiveness of mammography screening in decreasing breast cancer mortality among average-risk women aged 40 to 49 years and 70 years or older; the effectiveness of CBE and BSE in decreasing breast cancer mortality among women of any age; and the magnitude of harms of screening with mammography, CBE, and BSE.

**Methods**

The USPSTF and Agency for Healthcare Research and Quality (AHRQ) developed the key questions that guided our update. Investigators created an analytic framework incorporating the key questions and outlining the patient population, interventions, outcomes, and harms of the screening process (Appendix Figure 1, available at www.annals.org). The target population includes women without preexisting breast cancer and not considered to be at high risk for breast cancer on the basis of extensive family history of breast or ovarian cancer or other personal risk factors, such as abnormal breast pathology or deleterious genetic mutations. Harms include radiation exposure, pain during procedures, patient anxiety and other psychological responses, consequences of false-positive and false-negative test results, and overdiagnosis. “Overdiagnosis” refers to women receiving a diagnosis of invasive or noninvasive breast cancer who had abnormal lesions that were unlikely to become clinically evident during their lifetimes in the absence of screening (15). Overdiagnosis may have a greater effect on women with shorter life expectancies because of age or comorbid conditions.

**Data Sources and Searches**

We searched the Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews (through the fourth quarter of 2008) and MEDLINE (1 January 2001 to 1 December 2008) for relevant studies and meta-analyses (16). We also conducted secondary referencing by manually reviewing reference lists of key articles and searching citations by using Web of Science (17). Appendix Figure 2 (available at www.annals.org) shows our search results.

**Study Selection**

We selected studies on the basis of inclusion and exclusion criteria developed for each key question (16). To determine the effectiveness of screening, we included randomized, controlled trials (RCTs) and updates to previously published trials of screening with mammography (film and digital), MRI, CBE, or BSE with breast cancer mortality outcomes published since 2001. One trial was translated into English from Russian for this update (18). We also reviewed meta-analyses that included studies with mortality data. We excluded studies other than controlled trials and systematic reviews or those without breast cancer mortality as an outcome.

We determined harms of screening by using evidence from several study designs and data sources. For mammography, we focused our searches on recently published systematic reviews and meta-analyses of the harms previously described. We also conducted specific searches for primary studies published more recently than the included systematic reviews and meta-analyses. In addition, we evaluated data from the Breast Cancer Surveillance Consortium (BCSC), which is a collaborative network of 5 mammography registries and 2 affiliated sites with linkages to pathology and tumor registries across the United States that is sponsored by the National Cancer Institute (19, 20). These data draw from community samples that are representative of the larger, national population and may be more applicable to current practice in the United States than other published sources. Data include a mix of film and digital mammography. For harms of CBE and BSE, we reviewed screening trials of these procedures that reported potential adverse effects, used recently published systematic reviews, and conducted focused searches.

**Data Extraction and Quality Assessment**

We extracted details about the patient population, study design, analysis, follow-up, and results. By using predefined criteria developed by the USPSTF (21), 2 investigators rated the quality of each study as good, fair, or poor and resolved discrepancies by consensus. We included only systematic reviews rated as good quality in the report and RCTs rated as fair or good quality in the meta-analysis.

**Data Synthesis and Analysis**

**Meta-analysis of Mammography Trials**

We updated the 2002 meta-analysis to include new findings from published trials of mammography screening compared with control participants for women aged 40 to 49 years that reported relative risk (RR) reduction in breast cancer mortality. We conducted similar updates for other age groups for context. We used breast cancer mortality results from trials to estimate the pooled RR. We calculated estimates from a random-effects model under the Bayesian data analytic framework by using the RBugs package in R (22, 23), the same model as that used in the previous report (2). The Appendix (available at www.annals.org) provides additional details. We used
funnel plots to assess publication bias and L’Abbé plots to assess heterogeneity.

Analysis of BCSC Data

We obtained data from 600,830 women aged 40 years or older undergoing routine mammography screening from 2000 to 2005 at the BCSC sites from the BCSC Statistical Coordinating Center and stratified the data by age in decades. Routine screening was defined as having at least 1 mammogram within the previous 2 years, which is consistent with current USPSTF recommendations. For women who had several mammograms during the study, 1 result was randomly selected to be included in the calculations. These data constitute selected BCSC data intended to represent the experience of a cohort of regularly screened women without preexisting breast cancer or abnormal physical findings.

Variables include the numbers of positive and negative mammography results and, of these, the number of true-negative and false-negative results based on follow-up data within 1 year of mammography screening. A positive mammography result was defined according to standardized terminology and assessments of the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) manual used by the BCSC (24). These include 4 categories: needs additional evaluation (category 0), probably benign with a recommendation for immediate follow-up (category 3), suspicious (category 4), or highly suggestive of cancer (category 5) (25). For women who had a positive screening mammography result, additional data included the number of women undergoing additional imaging and biopsies; diagnoses, including invasive cancer and ductal carcinoma in situ; and negative results. We considered additional imaging procedures and biopsies done within 60 days of the screening mammography to be related to screening. From these data, we calculated age-specific rates (numbers per 1000 women per round) of invasive breast cancer, ductal carcinoma in situ, false-positive and false-negative mammography results, additional imaging, and biopsies. We based true-positive and true-negative mammography results on invasive and noninvasive cancer diagnosis. Rates of additional imaging and rates of biopsies may be underestimated because of incomplete capture of these examinations by the BCSC. The full evidence review (16) presents a sensitivity analysis of missing values; however, this does not include records that were unavailable to the BCSC.

Role of the Funding Source

The AHRQ funded this work, provided project oversight, developed key questions in conjunction with USPSTF members, and assisted with internal and external review of the draft manuscript but had no additional role in the design, conduct, or reporting of the review. Fifteen external experts not affiliated with the USPSTF reviewed the draft manuscript.

RESULTS

Breast Cancer Mortality Reduction With Mammography Screening for Women Aged 40 to 49 Years and 70 Years or Older (Key Question 1a)

The 2002 evidence review for the USPSTF included a meta-analysis (2) of 7 randomized trials of mammography screening rated as fair quality (26–28). Since then, a randomized trial from the United Kingdom that evaluated the effect of mammography screening, specifically in women aged 40 to 49 years, has been published (29), and data from a previously reported Swedish trial (30) have been updated. No trials of screening average-risk women that specifically evaluated the effectiveness of digital mammography or MRI have been published.

The Age trial (29) included 160,921 women aged 39 to 41 years who were randomly assigned from 1991 to 1997 to screening with annual mammography until 48 years of age or a control group who received usual care in the United Kingdom (Appendix Table 1, available at www.annals.org). After 10.7 years of follow-up, the RR was 0.97 (95% CI, 0.89 to 1.04) for all-cause mortality and 0.83 (CI, 0.66 to 1.04) for breast cancer mortality among women randomly assigned to screening. On the basis of the absolute reduction in breast cancer mortality among women randomly assigned to screening, the number needed to invite for screening to prevent 1 death from breast cancer over 10 years was 2512 (CI, 1149 to 13,544). The Age trial (29) met USPSTF criteria for fair rather than good quality because contamination of groups was not described and 70% or fewer women attended screening across the trial.

A new publication provides additional data from the Gothenburg trial (Appendix Table 1) (30). In this article, breast cancer mortality rates and risk ratios were calculated by using 3 methods, including a more comprehensive method that considers breast cancer mortality from cancer diagnosed during the follow-up phase of the trial. When this method was applied to women aged 39 to 49 years randomly assigned to screening at trial entry, the RR for breast cancer mortality was 0.69 (CI, 0.45 to 1.05) after 13 years of follow-up (30).

For women aged 39 to 49 years, 8 trials provided data for the meta-analysis, including 6 from the 2002 meta-analysis (Health Insurance Plan [HIP] of Greater New York [27], Canadian National Breast Screening Study-1 [CNBSS-1] [28], Stockholm [26], Malmö [26], Swedish Two-Country [2 trials] [26]), an update of the Gothenburg trial (30), and the Age trial (29). Combining results, the pooled RR for breast cancer mortality for women randomly assigned to mammography screening was 0.85 (95% credible interval [CrI], 0.75 to 0.96), which indicates a 15% reduction in breast cancer mortality in favor of screening (Figure). This corresponds to a number needed to invite for screening to prevent 1 breast cancer death of 1904 (CrI, 929 to 6378) over several screening rounds that varied by trial (2 to 9 rounds), and 11 to 20 years of follow-up. A funnel plot did not indicate the presence of publication bias, and a L’Abbé plot did
not reveal serious heterogeneity among the studies (16). Results are consistent with the 2002 meta-analysis (RR, 0.85 [CrI, 0.73 to 0.99]; 7 trials) (2, 3).

Sensitivity analysis excluded the HIP trial (27) because it was conducted more than 30 years ago and used outdated technology and the CNBSS-1 trial (28) because it enrolled prescreened volunteers rather than unselected samples. Exclusion of these trials did not significantly influence the results (16).

Results for women aged 70 years or older were confined to data from the Swedish Two-County trial (Östergötland) of women aged 70 to 74 years, precluding meta-analysis. These results indicate an RR for breast cancer mortality of 1.12 (CI, 0.73 to 1.72) (26), based on a more conservative determination of cause of death than previous reports (31, 32). The absolute numbers of deaths were not reported, the number of enrolled women was low (approximately 5000 in each group), and the number needed to screen was not estimable.

Meta-analyses of trials for women aged 50 to 59 years and 60 to 69 years were done to compare with results for women aged 40 to 49 years and 70 years or older (Table 1). Results are not directly similar to the 2002 meta-analysis that provided a combined estimate for women aged 50 to 59 years, yearly mammography screening for 12-view mammography averaging 7 mGy (33). For women aged 40 to 49 years, yearly mammography screening for 1

<table>
<thead>
<tr>
<th>Study/Author, Year (Reference)</th>
<th>Relative Risk for Breast Cancer Mortality (95% CrI)</th>
<th>Events/Total, n/n</th>
<th>Relative Risk for Breast Cancer Mortality (95% CrI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIP/Habbema et al, 1986 (27)</td>
<td>0.78 (0.56–1.08)</td>
<td>64/13 740</td>
<td>82/13 740</td>
</tr>
<tr>
<td>Kopperberg*/Tabar et al, 1995 (31)</td>
<td>0.72 (0.38–1.37)</td>
<td>22/9582</td>
<td>16/5031</td>
</tr>
<tr>
<td>CNBSS-1/Miller et al, 2002 (28)</td>
<td>0.97 (0.74–1.27)</td>
<td>105/25 214</td>
<td>108/25 216</td>
</tr>
<tr>
<td>Malmö/Nyström et al, 2002 (26)</td>
<td>0.73 (0.51–1.04)</td>
<td>53/13 568</td>
<td>66/12 279</td>
</tr>
<tr>
<td>Stockholm/Nyström et al, 2002 (26)</td>
<td>1.47 (0.77–2.78)</td>
<td>34/14 303</td>
<td>13/8021</td>
</tr>
<tr>
<td>Östergötland*/Nyström et al, 2002 (26)</td>
<td>1.05 (0.64–1.73)</td>
<td>31/10 285</td>
<td>30/10 459</td>
</tr>
<tr>
<td>Gothenburg/Bjurstam et al, 2003 (30)</td>
<td>0.70 (0.46–1.06)</td>
<td>34/11 724</td>
<td>59/14 217</td>
</tr>
<tr>
<td>Age/Moss et al, 2006 (29)</td>
<td>0.83 (0.66–1.04)</td>
<td>105/53 884</td>
<td>251/106 956</td>
</tr>
<tr>
<td>Total</td>
<td>0.85 (0.75–0.96)</td>
<td>448/152 300</td>
<td>625/195 919</td>
</tr>
</tbody>
</table>

CNBSS-1 = Canadian National Breast Screening Study-1; CrI = credible interval; HIP = Health Insurance Plan of Greater New York.

* Swedish Two-County trial.

Radiation Exposure

No studies directly measured the association between radiation exposure from mammography screening and breast cancer. Most x-rays are considered low-dose, low-energy radiation, with the mean glandular dose of bilateral, 2-view mammography averaging 7 mGy (33). For women aged 40 to 49 years, yearly mammography screening for 1

Table 1. Pooled RRs for Breast Cancer Mortality From Mammography Screening Trials for All Ages

<table>
<thead>
<tr>
<th>Age</th>
<th>Trials Included, n</th>
<th>RR for Breast Cancer Mortality (95% CrI)</th>
<th>NNI to Prevent 1 Breast Cancer Death (95% CrI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>39–49 y</td>
<td>8*</td>
<td>0.85 (0.75–0.96)</td>
<td>1904 (929–6378)</td>
</tr>
<tr>
<td>50–59 y</td>
<td>6†</td>
<td>0.86 (0.75–0.99)</td>
<td>1339 (322–7455)</td>
</tr>
<tr>
<td>60–69 y</td>
<td>2†</td>
<td>0.68 (0.54–0.87)</td>
<td>377 (230–1050)</td>
</tr>
<tr>
<td>70–74 y</td>
<td>1§</td>
<td>1.12 (0.73–1.72)</td>
<td>Not available</td>
</tr>
</tbody>
</table>

CrI = credible interval; NNI = number needed to invite to screening; RR = relative risk.

* Health Insurance Plan of Greater New York (27), Canadian National Breast Screening Study-1 (28), Stockholm (26), Malmö (26), Swedish Two-County trial (2 trials) (26, 31), Gothenburg trial (30), and Age trial (29).
† Canadian National Breast Screening Study-1 (28), Stockholm (26), Malmö (26), Swedish Two-County trial (2 trials) (26, 31), and Gothenburg trial (30).
‡ Malmö (26) and Swedish Two-County trial (Östergötland) (26).
§ Swedish Two-County trial (Östergötland) (26).
decade with potential additional imaging would expose an individual to approximately 60 mGy, although these levels vary (34). A recent systematic review included various types of studies of radiation exposure, such as radiation therapy, diagnostic radiation, and atomic bomb radiation, as the basis for predicting risk for inducing breast cancer (34). In studies of low-dose exposures, associations were inconsistent, whereas those of high-dose exposures indicated increased risk for breast cancer (34). The RRs in studies of high-dose exposures ranged from 1.33 to 11.39 for exposures of 0.3 to 43.4 Gy and were worse with higher doses of exposure, younger age at exposure, and longer follow-up (34). A more recent case–control study found that women exposed to diagnostic radiographs for screening or monitoring tuberculosis or pneumonia, or to therapeutic radiation for previous cancer, had increased risks for breast cancer (35).

Pain During Procedures

Breast compression is used during mammography to create uniform density, reduce breast thickness, and flatten overlying skin and tissues, which contributes to sharper images and reduces the radiation dose. However, compression may add to the discomfort of mammography for some women. A recent systematic review of 22 studies of pain and discomfort associated with mammography indicated that many women experience pain during the procedure (range, 1% to 77%), but few would consider this a deterrent from future screening (34). In these studies, pain was associated with the stage of the menstrual cycle, anxiety, and the anticipation of pain (34).

Anxiety, Distress, and Other Psychological Responses

Studies have shown conflicting results about anxiety, distress, and other psychological responses that result from mammography screening. A systematic review of 54 studies evaluated the adverse psychological effects of mammography screening programs (36). Most were cohort studies, and 24 used validated psychological measurement scales to assess the effects of screening. Studies indicated that women who received clear communication of their negative mammography results had minimal anxiety (36). Results were mixed in studies of women who were recalled for further testing as a result of screening. In several studies, women had persistent anxiety, despite eventual negative results, whereas some showed only transient anxiety (36). Some studies showed no differences between anxiety levels of women who had initial negative screening mammography results and those who had false-positive results (36).

A recent systematic review of 23 studies specifically examined the effects of false-positive screening mammography results on women aged 40 years or older (37). Twenty-six studies were included: 9 on psychological distress, 11 on anxiety, and 6 on worry. False-positive mammography results had no consistent effect on most women’s general anxiety and depression but increased breast cancer–specific distress, anxiety, apprehension, and perceived breast cancer risk for some (37).

False-Positive and False-Negative Mammography Results, Additional Imaging, and Biopsies

Published data on false-positive and false-negative mammography results, additional imaging, and biopsies that reflect current practices in the United States are limited. The probability of a false-positive screening mammography result was estimated at 0.9% to 6.5% in a meta-analysis of studies of sensitivity and specificity of mammography published 10 years ago (38). The cumulative risk for false-positive mammography results has been reported as 21% to 49% after 10 mammography examinations for women in general (39–41), and up to 56% for women aged 40 to 49 years (41). Additional data about mammography test performance indicate that sensitivity, recall rates, and cancer detection rates increase as the months since previous mammography increase, whereas specificity decreases (42). Few studies evaluate the effect of negative mammography results. Women stated that they would not delay evaluation of a new abnormal physical finding despite a previous negative mammography result in 1 survey (43).

Data from the BCSC for regularly screened women that are based on results from a single screening round indicate that false-positive mammography results are common in all age groups but are most common among women aged 40 to 49 years (97.8 per 1000 women per screening round) (Table 2). False-negative mammography results occur least among women aged 40 to 49 years (1.0 per 1000 women per screening round). Rates of additional imaging are highest among women aged 40 to 49 years (84.3 per 1000 women per screening round) and decrease with age, whereas biopsy rates are lowest among women aged 40 to 49 years (9.3 per 1000 women per screening round) and increase with age. The BCSC results indicate that for every case of invasive breast cancer detected by mammography screening in women aged 40 to 49 years, 556 women have mammography, 47 have additional imaging, and 5 have biopsies.

Overdiagnosis

A review of RCTs of mammography screening compared the cumulative incidence of breast cancer in intervention and control groups to determine the extent of overdiagnosis (44). In the 5 trials in which the control group did not receive screening, the absolute excess cumulative incidence of invasive and in situ breast cancer attributed to overdiagnosis among women randomly assigned to screening mammography ranged from 0.07 to 0.73 per 1000 woman-years.

Eight studies report estimates of overdiagnosis using different methods (16). Estimates are derived from data...
from screening programs in Italy (45), Denmark (46), and Norway and Sweden (47); a microsimulation model (48); analysis of incidence data from screening trials (46, 49, 50); and a Markov model with data from a screening trial (26) and several screening programs (51). None of these studies provide estimates specific to U.S. samples. Rates of overdiagnosis vary from less than 1% (45, 46, 49) to 30% (47), with most from 1% to 10%. Estimates differ by outcome (invasive vs. in situ breast cancer), by whether cases are incident or prevalent, and by age. The studies are too heterogeneous to combine statistically.

CBE Screening (Key Questions 1b and 2b)

Few trials have evaluated the effectiveness or harms of CBE in decreasing breast cancer mortality. In countries with widely practiced mammography screening, the use of CBE rests on its additional contribution to mortality reduction. The CNBSS-2 trial, which compares mammography with CBE versus CBE alone, showed no difference in mortality between these 2 approaches (52).

Three trials were designed to determine mortality outcomes by using CBE as the primary screening approach in countries with limited health care resources and without mammography screening programs (Appendix Table 2, available at www.annals.org). A randomized trial comparing CBE with no screening was conducted in the Philippines; however, it was discontinued after 1 screening round because of poor community acceptance and is inconclusive (53). Two randomized trials comparing CBE with no screening are ongoing in Egypt (54) and India (55).

In the pilot study for the Cairo Breast Screening Trial (54), 1.2% of women undergoing CBE had subsequent procedures with benign results. Of the 138,392 women examined in the Philippines study, 3479 had abnormal CBEs and 1220 completed diagnostic work-ups (53). Of these women, 34 (3%) had cancer, 563 (46%) had no detectable abnormalities, and 623 (51%) had biopsy results that were benign.

BSE (Key Questions 1c and 2c)

Preliminary results from trials of BSE in Russia and Shanghai were reviewed for the 2002 report (2), and final results have since been published (Appendix Table 2) (18, 56, 57). The effect of BSE on all-cause mortality in St. Petersburg, Russia, a community without routine mammography screening, was evaluated in a trial that met criteria for fair quality (18). Despite a significant increase in the number of cases of breast cancer detected when BSE instruction was provided, there was no reduction in all-cause mortality (RR, 1.07 [CI, 0.88 to 1.29]) (18). A good-quality randomized trial conducted in Shanghai, China, indicated breast cancer rates of 6.5 per 1000 for women instructed in BSE and 6.7 per 1000 for control participants after 11 years of follow-up (58). The number of women who died of breast cancer was the same in both groups (135 of 132,979 and 131 of 133,085, respectively; RR, 1.03 [CI, 0.81 to 1.31]). Published meta-analyses of randomized trials (59–61) and nonrandomized studies (59–61) of BSE also indicate no significant differences in breast cancer mortality between BSE and control groups.

In the Russian (18) and Shanghai (58) trials, more women randomly assigned to BSE had benign biopsy results than women in control groups (RR, 2.05 [CI, 1.80 to 2.33] for women in the Russian study and 1.57 [CI, 1.48 to 1.68] for women in the Shanghai study). A retrospective cohort study of 27,421 women aged 40 years or older in the United States indicated that those performing more frequent or longer-duration BSEs were more likely than...
<table>
<thead>
<tr>
<th>Table 3. Summary of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Studies and Type</td>
</tr>
<tr>
<td>Breast cancer mortality reduction with mammography screening (key question 1a)</td>
</tr>
<tr>
<td>Harms associated with mammography screening (key question 2a)</td>
</tr>
</tbody>
</table>

| Clinical breast examination screening benefits (key question 1b) | 1 (2 in progress) | RCT | The trial was discontinued after 1 round because of poor community acceptance. | Not applicable; poor | Poor | Inconclusive findings |

| Clinical breast examination screening harms (key question 2b) | 2 | 1 RCT and 1 descriptive study | Identified studies provide isolated descriptive data and are insufficient to address the question. | Not applicable; poor | Poor | Inconclusive findings |

| Breast self-examination screening benefits (key question 1c) | 2 trials and 3 systematic reviews | RCTs | Both trials were conducted in countries that do not have mass mammography screening. | Consistent; fair | Fair: Although trials were conducted in populations very different from the United States, results could be useful for U.S. practice. | Both trials indicated no reduction in mortality rates. |

Continued on following page
women with less frequent and shorter BSEs to have diagnostic mammography or ultrasonography (62). Contrary to the Russian and Shanghai studies, there was no significant association between BSE and biopsy rates in this study.

**DISCUSSION**

Table 3 summarizes the evidence for this review. Breast cancer mortality benefits from RCTs of screening are based on estimates of women who were randomly assigned to screening, whereas harms are based on data from women actually screened.

Trials of mammography screening for women aged 39 to 49 years indicate a statistically significant 15% reduction in breast cancer mortality for women randomly assigned to screening versus those assigned to controls. This translates to a number needed to invite for screening to prevent 1 breast cancer death of 1904 (CrI, 929 to 6378). These results are similar to those for women aged 50 to 59 years but less than those for women aged 60 to 69 years. For women aged 70 years or older, results from the Swedish Two-County trial (26) of women aged 70 to 74 years indicate no mortality reduction. However, these results are limited by including only a few women from 1 sample. Interpreting trial results stratified by age requires caution because except for the Age trial (29), age-specific results are subanalyses of trials designed for different purposes.

Although the addition of the Age trial (29) did not markedly change the results of the meta-analysis, its contribution to the evidence base is important. The Age trial (29) is the only trial of mammography that specifically evaluates the effectiveness of screening women in their 40s. It is the largest trial and draws from a community population. It is the most recent trial that reflects current screening, diagnostic, and treatment practices better than its predecessors, particularly those from the pretamoxifen era. As such, it is the most relevant trial. However, its results, although consistent with the meta-analysis in the direction of benefit, are not statistically significant. Also, its applicability to U.S. women is not clear, in light of important differences between mammography screening practices in the United States and the United Kingdom (63).

Harms of mammography screening have been identified, but their magnitude and effect are difficult to measure. The absolute level of radiation exposure and corresponding radiation risk from mammography is very low. Special considerations may be needed, however, for women exposed to additional radiation for other purposes or women particularly susceptible to radiation and breast cancer, such as BRCA mutation carriers. Patient adverse experiences, such as pain during procedures and anxiety and other psychological responses, are common but seem to be transient and do not adversely influence future screening practices. This may differ for individual women. Estimates of the magnitude of overdiagnosis vary depending on the analytic approach used. These estimates are difficult to apply because, for individual women, it is not known which types of cancer will progress, how quickly cancer will advance, and expected lifetimes.

The effectiveness of CBE has not been proven in large, well-designed trials. Current ongoing trials are limited to countries that do not provide routine mammography screening, which restricts their applicability to the United States. Work-ups for false-positive findings subject women to additional imaging and procedures countering the potential benefits of this low-technology approach. For BSE, the Russian (18) and Shanghai (58) trials simultaneously showed no reductions in mortality and increased numbers of benign biopsy results done as a result of BSE instruction.

Although more information is available to determine the benefits and harms of routine breast cancer screening in average-risk women, questions remain unanswered. The least amount of data is available for women aged 70 years or older, which is a rapidly growing population in the United States. Recent observational studies indicate that regular screening mammography among older women is associated with earlier-stage disease (64, 65) and lower breast cancer mortality rates (65). For the many older women who might live 20 to 30 years longer, breast cancer

---

**Table 3—Continued**

<table>
<thead>
<tr>
<th>Number of Studies and Type</th>
<th>Design</th>
<th>Limitations</th>
<th>Consistency; Overall Quality</th>
<th>Applicability</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast self-examination screening harms (key question 2c)</td>
<td>2 RCTs and 1 observational study</td>
<td>Both trials were conducted in countries that do not have mass mammography screening.</td>
<td>Not applicable; fair</td>
<td>Fair: Although trials were conducted in populations very different from the United States, results could be useful for U.S. practice.</td>
<td>2 trials indicated increased benign breast biopsies with breast self-examination instruction; biopsies were not increased in the observational study.</td>
</tr>
</tbody>
</table>

BCSC = Breast Cancer Surveillance Consortium; CrI = credible interval; MRI = magnetic resonance imaging; RCT = randomized, controlled trial.
detection and early treatment could reduce morbidity as well as mortality, thereby optimizing independence, function, quality of life, and costs of care in the final years.

Breast cancer is a continuum of entities, not just 1 disease that needs to be taken into account when considering screening and treatment options and when balancing benefits and harms. None of the screening trials consider breast cancer in this manner. As diagnostic and treatment experiences become more individualized (66) and include patient preferences, it becomes even more difficult to characterize benefits and harms in a general way.

New technologies, such as digital mammography and MRI, have become widely used in the United States without definitive studies of their effect on screening. Consumer expectations that new technology is better than old may obscure potential adverse effects, such as higher false-positive results and expense. No screening trials incorporating newer technology have been published, and estimates of benefits and harms in this report are based predominantly on studies of film mammography. No definitive studies of the appropriate interval for mammography screening exist, although trial data reflect screening intervals from 12 to 33 months.

Our meta-analysis of mammography screening trials indicates breast cancer mortality benefit for all age groups from 39 to 69 years, with insufficient data for older women. False-positive results are common in all age groups and lead to additional imaging and biopsies. Women aged 40 to 49 years experience the highest rate of additional imaging, whereas their biopsy rate is lower than that for older women. Mammography screening at any age is a tradeoff of a continuum of benefits and harms. The ages at which this tradeoff becomes acceptable to individuals and society are not clearly resolved by the available evidence.

From Oregon Health & Science University; Veterans Affairs Medical Center; and the Women and Children’s Health Research Center, Providence Health & Services, Portland, Oregon.

Disclaimer: The findings and conclusions in this document are those of the authors, who are responsible for its content, and do not necessarily represent the views of the AHRQ. No statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

Acknowledgment: The authors acknowledge the contributions of AHRQ Project Officer Mary Barton, MD, MPH, and USPSTF leads Russ Harris, MD, MPH; Allen Dietrich, MD; Carol Loveland-Cherry, PhD, RN; Judith Ockene, PhD, Med; and Berndette Melnyk, PhD, RN, CPNP/NPP. The authors thank the BCSC investigators, participating mammography facilities, and radiologists for the data used in this project. A list of the BCSC investigators and procedures for requesting BCSC data for research purposes are available at http://breastscreening.cancer.gov/. They also thank Patricia A. Carney, PhD; Steve Taplin, MD; Sebastien Haneuse, PhD; and Rod Walker, MS; for direct work with investigators, along with Andrew Hamilton, MLS, MS, for literature searches, and Sarah Baird, MS, for managing the bibliography at the Oregon Evidence-based Practice Center at the Oregon Health & Science University.
Screening for Breast Cancer


**APPENDIX: DETAILS OF THE META-ANALYSIS**

The meta-analysis is an update of the previous 2002 meta-analysis that includes results from published trials of mammography screening for women aged 40 to 49 years that report reduction in breast cancer mortality. With the addition of only 1 new data point, the meta-analysis for the update was less extensive than the 2002 meta-analysis. We did not update the model for RR and length of follow-up (the 2-level hierarchical model). We conducted similar updates for other age groups for context.

As with the original 2002 meta-analysis, we estimated the model by using a Bayesian data analytic framework, but this time using the BRugs package in R (22, 23). BRugs is an R interface to OpenBUGS, the successor to WinBUGS. The R code to create the data set is below.

```r
# R code to create dataset
doctordataset <- c('Age', 'CNBSS-1', 'HIP', 'Gothenburg', 'Stockholm', 'Malmo', 'Kopparberg', 'Ostergotland')
y.int <- c(105, 105, 64, 34, 34, 53, 22, 31) n.int <- c(53884, 25214, 13740, 11724, 14303, 13568, 59582, 10285) py.int <- c(578390, 282606, 192360, NA, 203000, 184000, 124566, 172000) y.cntl <- c(251, 108, 82, 59, 13, 66, 16, 30) n.cntl <- c(106956, 25216, 13740, 14217, 8021, 12279, 16000, 65403, 67600) py.cntl <- c(1149380, 282575, 192360, NA, 117000, 160000, 65403, 176000) n <- 10000
rate.int <- n * y.int / n.int rate.cntl <- n * y.cntl / n.cntl rr <- rate.int / rate.cntl rd <- rate.int - rate.cntl nns <- 1 / ((y.cntl/n.cntl) - (y.int / n.int)) dataset <- data.frame( study, y.int, n.int, py.int, rate.int, y.cntl, n.cntl, py.cntl, rate.cntl, rr, rd, nns )
# Save dataset for BRugs to use
dataset.bugs <- cbind(y.int, n.int, y.cntl, n.cntl) colnames(dataset.bugs) <- c("y.int", "n.int", "y.cntl", "n.cntl") bugsData(data.frame(dataset.bugs), fileName="dataset.bugs", digits = 5) constants <- cbind(nrow(dataset.bugs)) colnames(constants) <- c("n") bugsData(data.frame(constants), fileName="constants.bugs", digits = 1)

The model assumes that the number of deaths from each study come from a binomial distribution with the probability parameter of \( \alpha \) for the control group and \( \alpha + \beta \) for the screening group. A random component, \( \sigma_z \), is added to both probability parameters to allow for the random effect of the study \( i \). Non-informative prior probability distributions were used.

```
The convergence of the parameter estimation was assessed and deemed adequate from the 10 000 burn-in draws. Next, we generated 100 000 draws from the 4 chains. These draws were thinned to yield a sample of 1000 uncorrelated estimates from the posterior distributions.

After the model was estimated and the samples were thinned, point estimates (mean) and 95% CrIs (2.5 and 97.5 percentiles) for RR, risk difference, and number needed to invite to screening were calculated.

- The pooled number needed to invite to screening could be misleading if the baseline risk for mortality is appreciably varied between studies (67). One recommendation to accommodate this is to apply the pooled RR estimate to a range of control rates and then calculate the number needed to invite to screening. The pooled rate of mortality among the control groups of our studies was estimated to be 35.5 deaths per 10 000 women (95% CrI, 25.1 to 48.3). The range of mortality rates among the control groups was 16.2 to 59.7 per 10 000 women. Applying the pooled RR estimate of 0.85 to the high end of the mortality rate range (59.7) yields a number needed to invite to screening estimate of 1116 per 10 000 women. Applying the pooled RR estimate of 0.85 to the low end of the mortality rate range (16.2) yields a number needed to invite to screening estimate of 4115 per 10 000 women. This range 1116 to 4115 per 10 000 women is between studies (67). One recommendation to accommodate this is to apply the pooled RR estimate to a range of control rates and then calculate the number needed to invite to screening. The pooled rate of mortality among the control groups of our studies was estimated to be 35.5 deaths per 10 000 women (95% CrI, 25.1 to 48.3). The range of mortality rates among the control groups was 16.2 to 59.7 per 10 000 women. Applying the pooled RR estimate of 0.85 to the high end of the mortality rate range (59.7) yields a number needed to invite to screening estimate of 1116 per 10 000 women. Applying the pooled RR estimate of 0.85 to the low end of the mortality rate range (16.2) yields a number needed to invite to screening estimate of 4115 per 10 000 women. This range 1116 to 4115 per 10 000 women is within the 95% CrI that we report for number needed to invite to screening that we estimated from the posterior distributions of our mortality rate estimates. Alternatively, the bounds of our 95% CrI to number needed to invite to screening correspond to a range of control group mortality rates of 10.5 to 71.8 per 10 000 women, a range beyond that seen in the studies included in our analysis.
Appendix Figure 1. Analytic framework and key questions.

Key Questions
1a. Does screening with mammography (film and digital) or MRI decrease breast cancer mortality among women aged 40 to 49 y and ≥70 y?
1b. Does CBE screening decrease breast cancer mortality? Alone or with mammography?
1c. Does BSE practice decrease breast cancer mortality?
2a. What are the harms associated with screening with mammography (film and digital) and MRI?
2b. What are the harms associated with CBE?
2c. What are the harms associated with BSE?

BSE = breast self-examination; CBE = clinical breast examination; MRI = magnetic resonance imaging.
* Includes radiation exposure, pain, psychological responses, false-positive and false-negative test results, and overdiagnosis.
Appendix Figure 2. Literature search and selection.

Abstracts of potentially relevant articles identified through MEDLINE, Cochrane*, Web of Science, and other sources† (n = 2994)

Excluded abstracts and background articles (n = 2435)

Full-text articles reviewed for relevance to key questions (n = 559)

Excluded articles (n = 514)
- Wrong population (including high-risk): 20
- Wrong intervention: 4
- Wrong outcome: 39
- Wrong study design or no original data for meta-analysis: 160
- Development of technology: 13
- Does not address a key question: 80
- Treatment-focused: 43
- Wrong age: 2
- Does not break out data by age for meta-analysis: 3
- Contextual only: 104
- Non-English-language: 3
- Covered by included papers or previous USPSTF report: 43

Included articles‡

Key question 1a: mammography outcomes

Age 40–49 y: 1 new trial + updated data from 1 prior trial + 7 trials from prior review
Age 70–74 y: 1 trial from prior review

Key question 1b: CBE outcomes

1 new trial + 1 trial from prior review + 2 unfinished trials

Key question 1c: BSE outcomes

2 trials + 3 SRs

Key question 2a: mammography harms

Radiation: 5 studies + 1 SR
Pain during procedures: 2 SRs
False-positive results: 3 studies + 2 SRs
False-negative results: 3 studies
Overdiagnosis: 9 studies + 1 SR
Anxiety, distress: 2 SRs
Personal cost: 1 study

Key question 2b: CBE harms

3 studies

Key question 2c: BSE harms

3 studies

Excluded abstracts and background articles (n = 2435)

Abstracts of potentially relevant articles identified through MEDLINE, Cochrane*, Web of Science, and other sources† (n = 2994)

Included articles‡

BSE = breast self-examination; CBE = clinical breast examination; SR = systematic review; USPSTF = U.S. Preventive Services Task Force.
* Cochrane databases include the Cochrane Central Register of Controlled Trials and the Cochrane Database of Systematic Reviews.
† Other sources include reference lists and studies suggested by experts.
‡ Some articles are included for more than 1 key question.
### Appendix Table 1. Mammography Screening Trials Included in Meta-analysis

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Baseline Study Year</th>
<th>Setting or Population (Screened Patients; Control Participants)</th>
<th>Enrollment Age, y</th>
<th>Randomization Method</th>
<th>Study Group</th>
<th>Screening Protocol</th>
<th>Follow-up, y</th>
<th>USPSTF Quality Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Insurance Plan of Greater New York, 1986 (27)</td>
<td>1963</td>
<td>New York health plan members (30,239; 30,256)</td>
<td>40–64</td>
<td>Pairs of women stratified by age and family size were individually randomly assigned by a drawing from a list</td>
<td>Mammography + CBE vs. usual care</td>
<td>12</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Canadian National Breast Screening Study-1, 2002 (28)</td>
<td>1980</td>
<td>15 centers in Canada, self-selected participants (25,214; 25,216)</td>
<td>40–49</td>
<td>Blocks were stratified by center and 5-y age group after CBE</td>
<td>Mammography + CBE vs. usual care (all women prescreened and instructed in BSE)</td>
<td>12</td>
<td>4–5</td>
<td>2</td>
</tr>
<tr>
<td>Gothenburg Breast Screening trial, 2003 (30)*</td>
<td>1982</td>
<td>All women born from 1923–1944 who lived in Gothenburg, Sweden (20,724; 28,809)</td>
<td>39–59</td>
<td>Cluster, based on day of birth (1923–1935 cohort [18%]) and individual (1936–1944 cohort [82%])</td>
<td>Mammography vs. usual care; control participants offered screening after 5 y and completed screening at approximately 7 y</td>
<td>18</td>
<td>5</td>
<td>1–2</td>
</tr>
<tr>
<td>Stockholm, 2002 (26)</td>
<td>1981</td>
<td>Residents of southeast greater Stockholm, Sweden (40,318; 19,943)</td>
<td>40–64</td>
<td>Individual, by day of month; screening to control group ratio is 2:1</td>
<td>Mammography vs. usual care</td>
<td>24–28</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Malmö, 2002 (26)</td>
<td>1976–1978</td>
<td>All women born from 1927–1945 living in Malmö, Sweden (21,088; 21,195)</td>
<td>45–70</td>
<td>Individual, within birth year</td>
<td>Mammography vs. usual care; control participants offered screening after 14 y</td>
<td>18–24</td>
<td>9</td>
<td>1–2</td>
</tr>
<tr>
<td>Swedish Two-County trial (2 trials), 2002 (26); 1995 (31)</td>
<td>1977</td>
<td>From Östergötland and Kopparberg counties in Sweden (77,080; 55,985)</td>
<td>40–74</td>
<td>Clusters, based on geographic units; blocks designed to be demographically homogeneous</td>
<td>Mammography vs. usual care; control participants offered screening after 7 y</td>
<td>24–33</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

BSE = breast self-examination; CBE = clinical breast examination; USPSTF = U.S. Preventive Services Task Force.

* New data since previous recommendation.
### Appendix Table 2. Trials of CBE and BSE

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Technique</th>
<th>Years</th>
<th>Setting or Population (Screened Patients; Control Participants)</th>
<th>Enrollment Age, y</th>
<th>Study Design</th>
<th>Intervention</th>
<th>Outcomes and Ratings</th>
<th>USPSTF Quality Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pisani et al, 2006 (53)</td>
<td>CBE</td>
<td>1996–1997</td>
<td>Manila, Philippines; women living in the 12 central areas (151 168; control participants not indicated)</td>
<td>35–64</td>
<td>RCT; block randomization of 202 health centers</td>
<td>5 annual CBEs vs. usual care provided by nurses and midwives; CBE instruction using the Mammmacare program*</td>
<td>Breast cancer mortality not reported</td>
<td>Poor: low participation; discontinued after 1 round</td>
</tr>
<tr>
<td>Boulos et al, 2005 (54)</td>
<td>CBE or BSE</td>
<td>Pilot: 2000–2002; RCT: ongoing</td>
<td>Cairo, Egypt; women living in area around Italian Hospital (1924; 1927)</td>
<td>39–65</td>
<td>RCT; block randomization</td>
<td>CBE/BSE twice (intervention) vs. CBE/BSE once (control) provided by female physicians; CBE training at Italian Hospital 2 mo before study</td>
<td>Breast cancer incidence</td>
<td>Benign procedures: 1.2% after 1 round</td>
</tr>
<tr>
<td>National Cancer Institute (55)†</td>
<td>CBE or BSE</td>
<td>1998 and ongoing</td>
<td>Mumbai, India; women living in area around Tata Memorial Hospital (150 000; control participants not indicated)</td>
<td>35–64</td>
<td>RCT; cluster randomization</td>
<td>CBE + BSE + breast health education every 24 mo for 4 rounds vs. education alone provided by trained female health workers; CBE training for 5 mo before trial</td>
<td>Breast cancer mortality</td>
<td>Not available Not rated (in progress)</td>
</tr>
<tr>
<td>Thomas et al, 2002 (58)</td>
<td>BSE</td>
<td>1989–2000</td>
<td>Shanghai, China; women working at 1 of 519 factories (132 979; 133 085)</td>
<td>31–65</td>
<td>RCT; factories assigned to BSE or control group</td>
<td>BSE instruction with periodic reinforcement provided by trained former factory medical workers vs. no instruction; initial BSE instruction, follow-up sessions at 1 and 3 y, medically supervised BSE every 6 mo</td>
<td>Breast cancer mortality: RR, 1.03 (95% CI, 0.81–1.31)</td>
<td>Benign biopsies: RR, 1.57 (CI, 1.48–1.68) Good</td>
</tr>
<tr>
<td>Semiglazov et al, 2003 (18)</td>
<td>BSE</td>
<td>1985–2001</td>
<td>St. Petersburg, Russia; women attending 1 of 28 clinics (58 985; 64 763)</td>
<td>40–64</td>
<td>RCT; cluster randomization</td>
<td>BSE instruction with refresher every 3 y provided by trained nurses or physicians vs. no instruction; providers received 3-h training; instruction given to groups of 5–20 women</td>
<td>All-cause mortality: RR, 1.07 (CI, 0.88–1.29)</td>
<td>Benign biopsies: RR, 2.05 (CI, 1.80–2.33) Fair: low adherence; inconsistent data reported</td>
</tr>
</tbody>
</table>

BSE = breast self-examination; CBE = clinical breast examination; RCT = randomized, controlled trial; RR = relative risk; USPSTF = U.S. Preventive Services Task Force.

* Gainesville, Florida.
† Risks are not calculated because diagnostic follow-up for a positive CBE was 35%.
‡ Trial is in progress.