In 2009, the U.S. Preventive Services Task Force (USPSTF) recommended biennial mammography screening for women aged 50 to 74 years (1) on the basis of evidence of benefits and harms (2). The USPSTF concluded that screening decisions for women aged 40 to 49 years should be based on individual considerations, and that evidence was insufficient to assess benefits and harms for women aged 75 years or older (1).

Mammography screening in the United States is generally opportunistic, unlike many screening programs organized as public health services in other countries. Despite changes in practice guidelines and variation in clinical practices (3), overall screening rates in the United States have remained relatively stable for the past decade (4, 5). Data from the Healthcare Effectiveness Data and Information Set indicate that mammography screening in 2014 in HMOs was performed for 74% of eligible women covered by commercial plans, 72% by Medicare, and 59% by Medicaid (6).

This systematic review updates evidence for the USPSTF on the effectiveness of mammography screening in reducing breast cancer mortality, all-cause mortality, and advanced breast cancer for women at average risk; and how effectiveness varies by age, risk factors, screening intervals, and imaging modalities. Systematic reviews of harms of screening (7), performance characteristics of screening methods (8), and accuracy of breast density determination and use of supplemental screening technologies (9) are provided in separate reports.

**Methods**

**Scope, Key Questions, and Analytic Framework**

The USPSTF determined the scope and key questions for this review by using established methods (10, 11). A standard protocol was developed and publicly posted on the USPSTF Web site. A technical report fur-
ther describes the methods and includes search strategies and additional information (7).

Investigators created an analytic framework outlining the key questions, patient populations, interventions, and outcomes reviewed (Appendix Figure 1, available at www.annals.org). Key questions include the effectiveness of screening in reducing breast cancer mortality, all-cause mortality, and advanced breast cancer, and how effectiveness differs by age, risk factors, screening intervals, and modalities (mammography [film, digital, tomosynthesis], magnetic resonance imaging [MRI], and ultrasonography).

The target population for the USPSTF recommendation includes women aged 40 years or older, and excludes women with known physical signs or symptoms of breast abnormalities and those at high-risk for breast cancer whose surveillance and management are beyond the scope of the USPSTF’s recommendations for prevention services (i.e., preexisting breast cancer or high-risk breast lesions, hereditary genetic syndromes associated with breast cancer, or previous large doses of chest radiation before age 30 years). Risk factors considered in this review are common among women who are not at high risk for breast cancer (12) (Appendix Figure 1, available at www.annals.org).

Data Sources and Searches
A research librarian conducted electronic database searches of the Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and Ovid MEDLINE to 4 June 2015. Searches were supplemented by references identified from additional sources, including reference lists and experts. Additional unpublished data were provided by the investigators of the Canadian National Breast Screening Study (CNBSS) and Swedish Two-County Trial.

Study Selection
Two investigators independently evaluated each study to determine inclusion eligibility on the basis of prespecified criteria. Discrepancies were resolved through consensus.

We included randomized, controlled trials (RCTs); observational studies of screening cohorts; and systematic reviews that compared outcomes of women exposed to screening versus not screening. For advanced cancer outcomes, studies that reported the incidence of late-stage disease among screened and unscreened populations were included, whereas those reporting comparisons of detection methods that did not capture a woman’s longitudinal screening experience were not included (e.g., rates of screen-detected vs. non-screen-detected cancer).

Studies providing outcomes specific to age, risk factors, screening intervals, and modalities were preferred over studies providing general outcomes, when available. Studies most clinically relevant to practice in the United States were selected over studies that were less relevant. Relevance was determined by practice setting, population, date of publication, and use of technologies and therapies in current practice. Studies meeting criteria for high quality and those with designs ranked higher in the study design-based hierarchy of evidence were emphasized because they are less susceptible to bias (e.g., RCTs over observational studies).

Data Extraction and Quality Assessment
Details of the study design, patient population, setting, screening method, interventions, analysis, follow-up, and results were abstracted by one investigator and confirmed by a second. Two investigators independently applied criteria developed by the USPSTF (10, 11) to rate the quality of each study as good, fair, or poor for studies designed as RCTs, cohort studies, case-control studies, and systematic reviews; criteria to rate other study designs included in this review are not available. Discrepancies were resolved through consensus.

Data Synthesis
We conducted several meta-analyses to determine more precise summary estimates when adequate data were reported by trials rated as fair- or good-quality. In each meta-analysis, the number of included trials was counted as the number of discrete data sources contributing to the summary estimate using their most recent results. To determine the appropriateness of meta-analysis, we considered clinical and methodological diversity and assessed statistical heterogeneity. All outcomes were binary (breast cancer mortality, all-cause mortality, and advanced cancer incidence defined by stage and tumor size). We used a random-effects model to combine relative risks (RRs) as the effect measure of the meta-analyses, while incorporating variation among studies. A profile-likelihood model was used to combine studies in the primary analyses (13). We assessed the presence of statistical heterogeneity among the studies by using the standard Cochran chi-square test, and the magnitude of heterogeneity by using the $I^2$ statistic (14).

To account for clinical heterogeneity and obtain clinically meaningful estimates, we stratified the analyses by age group whenever possible (39 to 49 years, 50 to 59 years, 60 to 69 years, 70 to 74 years, and ≥50 years). We obtained additional age-stratified data for the meta-analysis from the investigators of 3 trials (15, 16) (Tabár L. Personal communication).

For breast cancer mortality, we used 2 methods of including cases to help clarify discrepancies between estimates. The long case accrual method counts all breast cancer cases contributing to breast cancer deaths. In this method, the case accrual time is equivalent to or close to the follow-up time. The short case accrual method includes only deaths that occur among cases of breast cancer diagnosed during the screening intervention period, and in some trials, within an additional defined case accrual period. The longest follow-up times available for each trial were selected for inclusion in the initial meta-analyses, and sensitivity analyses were conducted by using results of short case accrual methods.

We calculated the absolute rate reduction for 100 000 woman-years of follow-up (i.e., 10 000 women
followed for 10 years) for each age group on the basis of the combined RR and the combined cancer rate of the control group. We estimated combined cancer rates for each age group for controls with a random effects Poisson model using data from the trials. All analyses were performed by using Stata/IC, version 13.1 (StataCorp).

We assessed the aggregate internal validity (quality) of the body of evidence for each key question as good, fair, or poor by using methods developed by the USPSTF that are based on the number, quality, and size of studies; consistency of results between studies; and directness of evidence (10, 11).

Role of the Funding Source

This research was funded by the Agency for Healthcare Research and Quality (AHRQ) under a contract to support the work of the USPSTF. The investigators worked with USPSTF members and AHRQ staff to develop and refine the scope, analytic framework, and key questions; resolve issues during the project; and finalize the report. The AHRQ had no role in study selection, quality assessment, synthesis, or development of conclusions. The AHRQ provided project oversight; reviewed the draft report; and distributed the draft for peer review, including to representatives of professional societies and federal agencies. The AHRQ performed a final review of the manuscript to ensure that the analysis met methodological standards. The investigators are solely responsible for the content and the decision to submit the manuscript for publication.

RESULTS

Of the 12 070 abstracts identified by electronic searches and other sources, 38 studies met inclusion criteria for key questions in this report, including 5 systematic reviews of 62 studies (Appendix Figure 2, available at www.annals.org).

RCTs of Screening

Eight main trials of mammography screening met criteria for fair quality: the Health Insurance Plan of Greater New York (HIP) trial (17); the CNBSS (Canadian National Breast Cancer Screening)-1 (18, 19) and CNBSS-2 (20, 21); the Age trial, performed in the United Kingdom (22); and 4 trials from Sweden, which were the Stockholm trial (23), Malmö Mammographic Screening Trial (referred to separately as “MMST I” and “MMST II”) (24), Gothenburg trial (25), and Swedish Two-County Trial (referred to separately as “Östergötland” and “Kopparberg”) (26). Updates of the CNBSSs, the Age trial, and Swedish Two-County Trial provided new data for this report (15, 16, 27). The Edinburgh trial (28) was not included because of important baseline differences between screening and control groups, suggesting inadequate randomization.

Trials included over 600 000 women and varied in their recruitment, randomization, screening protocols, control groups, and sizes (Appendix Table 1, available at www.annals.org). Breast cancer mortality was the main outcome measure, and all trials evaluated differences between screening and control groups on an intention-to-screen basis. Other important characteristics are described in the technical report (7).

Breast Cancer Mortality Outcomes

Screening Trials

The Swedish Two-County Trial (Kopparberg and Östergötland [26]), Age trial (27), Gothenburg trial (25), and CNBSS-1 and CNBSS-2 (15) used long case accrual methods to report breast cancer mortality by age. The HIP trial (29), MMST I, MMST II, and Stockholm trials used only short case accrual (30) to report breast cancer mortality by age. Across all trials, the mean or median screening intervention time ranged from 3.5 to 14.6 years, case accrual time from 7.0 to 17.4 years, and follow-up time from 11.2 to 21.9 years.

For women aged 39 to 49 years, the combined RR for breast cancer mortality was 0.92 (95% CI, 0.75 to 1.02) (9 trials [15, 25–27, 29, 30]) (Figure 1). The combined RR for women aged 50 to 59 years was 0.86 (CI, 0.68 to 0.97) (7 trials [15, 25, 26, 29, 30]); for those aged 60 to 69 years, it was 0.67 (CI, 0.54 to 0.83) (5 trials [26, 29, 30]). Combining results for women aged 50 to 69 years indicated an RR of 0.78 (CI, 0.68 to 0.90). Only 3 Swedish trials (Östergötland, Kopparberg, and MMST I) provided outcomes for women aged 70 to 74 years, and the numbers of events in these trials were much lower than for other age groups (26). A meta-analysis indicated a combined RR of 0.80 (CI, 0.51 to 1.28) (3 trials [26]).

Results of the meta-analysis were used to determine absolute rates of breast cancer mortality reduction per 10 000 women screened for 10 years (Table 1). The number of deaths reduced (i.e., prevented) was 2.9 (CI, 0.6 to 8.9) for women aged 39 to 49 years; 7.7 (CI, 1.6 to 17.2) for those aged 50 to 59 years; 21.3 (CI, 10.7 to 31.7) for those aged 60 to 69 years; and 12.5 (CI, –17.2 to 32.1) for those aged 70 to 74 years. Absolute reduction for the combined group of women aged 50 to 69 years was 12.5 (CI, 5.9 to 19.5).

The effect of screening was diminished, although the statistical significance of the estimates did not change, in our sensitivity analysis that included results of a published combined analysis of the Swedish trials (MMST I, MMST II, Stockholm, Östergötland, Gothenburg, and Stockholm) using a long case accrual (“follow-up”) method (30) rather than results of individual trials. In a separate sensitivity analysis, meta-analysis estimates from trials with short case accrual methods differed only slightly from those with long case accrual (Table 1). Across all trials with short case accrual, the mean or median screening intervention time ranged from 3.5 to 14.6 years, case accrual time from 5.0 to 15.5 years, and follow-up time from 10.7 to 25.7 years.

Observational Studies

Three good-quality systematic reviews of observational studies of screening were recently conducted by the EUROSCREEN Working Group to assess the effectiveness of mammography screening on breast cancer mortality (31–33). An additional review included many
of the same studies (34). The EUROSCREEN reviews included studies from current population-based screening programs in Europe and the United Kingdom; included women aged 50 to 69 years; and were designed as time–trend, incidence-based mortality, or case–control studies. Although quality criteria were not prespecified, the studies were subjected to critical review according to design-specific factors.

Of 12 time-trend studies reporting changes in breast cancer mortality in relation to the introduction of screening (32, 35–39, 40–46), 3 with adequate follow-up reported mortality reductions ranging from 28% to 35% (41, 42, 45). A meta-analysis (33) of incidence-based mortality studies estimating breast cancer mortality from a cohort of women not invited for screening, or from historical and current control

---

**Figure 1.** Effects of screening on breast cancer mortality.

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Trial Name</th>
<th>Mean Follow-up, y</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST II</td>
<td>11.2</td>
<td>0.64 (0.39–1.06)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.73 (0.37–1.41)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Östergötland</td>
<td>12.5</td>
<td>1.02 (0.52–1.99)</td>
</tr>
<tr>
<td>Moss et al, 2015 (27)</td>
<td>Age</td>
<td>17.5</td>
<td>0.93 (0.80–1.09)</td>
</tr>
<tr>
<td>Bjerstam et al, 2003 (25)</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.69 (0.45–1.05)</td>
</tr>
<tr>
<td>Habbema et al, 1986 (29)</td>
<td>HIP</td>
<td>14.0</td>
<td>0.75 (0.53–1.05)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>Stockholm</td>
<td>14.3</td>
<td>1.52 (0.80–2.88)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST I</td>
<td>18.2</td>
<td>0.74 (0.42–1.29)</td>
</tr>
<tr>
<td>Miller et al, 2014 (15)</td>
<td>CNBSS-1</td>
<td>21.9</td>
<td>1.04 (0.87–1.24)</td>
</tr>
<tr>
<td>Overall ($I^2 = 25%$; $P = 0.230$)</td>
<td></td>
<td></td>
<td>0.92 (0.75–1.02)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Trial Name</th>
<th>Mean Follow-up, y</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Östergötland</td>
<td>12.5</td>
<td>0.85 (0.52–1.38)</td>
</tr>
<tr>
<td>Nystrom et al, 2002 (30)*</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.48 (0.29–0.77)</td>
</tr>
<tr>
<td>Bjerstam et al, 2003 (25)</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.83 (0.60–1.15)</td>
</tr>
<tr>
<td>Habbema et al, 1986 (29)</td>
<td>HIP</td>
<td>14.0</td>
<td>0.83 (0.61–1.13)</td>
</tr>
<tr>
<td>Nystrom et al, 2002 (30)*</td>
<td>MMST I</td>
<td>18.1</td>
<td>0.98 (0.75–1.29)</td>
</tr>
<tr>
<td>Miller et al, 2014 (15)</td>
<td>CNBSS-2</td>
<td>21.9</td>
<td>0.94 (0.78–1.13)</td>
</tr>
<tr>
<td>Overall ($I^2 = 38.0%$; $P = 0.139$)</td>
<td></td>
<td></td>
<td>0.86 (0.68–0.97)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Trial Name</th>
<th>Mean Follow-up, y</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.58 (0.35–0.96)</td>
</tr>
<tr>
<td>Nystrom et al, 2002 (30)*</td>
<td>Östergötland</td>
<td>12.5</td>
<td>0.62 (0.43–0.91)</td>
</tr>
<tr>
<td>Habbema et al, 1986 (29)</td>
<td>HIP</td>
<td>14.0</td>
<td>0.85 (0.48–1.47)</td>
</tr>
<tr>
<td>Nystrom et al, 2002 (30)*</td>
<td>MMST I</td>
<td>15.5</td>
<td>0.64 (0.45–0.92)</td>
</tr>
<tr>
<td>Overall ($I^2 = 0.0%$; $P = 0.739$)</td>
<td></td>
<td></td>
<td>0.67 (0.54–0.83)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Trial Name</th>
<th>Mean Follow-up, y</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Östergötland</td>
<td>12.5</td>
<td>0.82 (0.43–1.58)</td>
</tr>
<tr>
<td>Nystrom et al, 2002 (30)*</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.76 (0.42–1.36)</td>
</tr>
<tr>
<td>Nystrom et al, 2002 (30)*</td>
<td>MMST I</td>
<td>13.6</td>
<td>0.98 (0.15–6.60)</td>
</tr>
<tr>
<td>Overall ($I^2 = 0.0%$; $P = 0.962$)</td>
<td></td>
<td></td>
<td>0.80 (0.51–1.28)</td>
</tr>
</tbody>
</table>

Meta-analysis of trials using the longest follow-up times available. CNBSS = Canadian National Breast Screening Study; HIP = Health Insurance Plan of New York; MMST = Malmö Mammographic Screening Trial.

* Used short case accrual.
groups, indicated a risk reduction of 0.75 (CI, 0.69 to 0.81) (7 studies [42, 47–52]) for women invited to screening and 0.62 (CI, 0.56 to 0.69) (7 studies [42, 47–52]) for those actually screened. A meta-analysis of case-control studies (33) indicated an odds ratio of 0.69 (CI, 0.57 to 0.83) (7 studies [53–58]) for women invited to screening; and 0.52 (CI, 0.42 to 0.65) (7 studies [53–58]) for those actually screened.

Six additional studies were not included in the published systematic reviews because they were published in 2011 or later (59–63), included women in countries outside Europe and the United Kingdom (63, 64), or focused on ages older or younger than 50 to 69 years (59, 64) (Appendix Table 2, available at www.annals.org). These studies are generally consistent with the EUROSCREEN results (61–63), including 2 observational studies of women in their 40s indicating a 26% to 44% reduction in breast cancer mortality for women invited to (59) or participating in screening (59, 63).

### All-Cause Mortality

All included trials of mammography screening reported all-cause mortality outcomes. However, not all trials reported them according to age, and the 2 Canadian trials reported combined estimates. For all ages, the combined RR of 0.99 (CI, 0.97 to 1.002) (9 trials [15, 27, 30, 65]) was consistent with no reduction in all-cause mortality with screening. Results were similar for each age group (RR, 0.99 [CI, 0.94 to 1.05] for women aged 39 to 49 years [7 trials], 1.02 [CI, 0.94 to 1.10] for those aged 50 to 59 years [3 trials], 0.97 [CI, 0.90 to 1.04] for those aged 60 to 69 years [2 trials], and 0.98 [CI, 0.86 to 1.14] for those aged 70 to 74 years [2 trials]) and when short case accrual methods were used.

### Mortality Outcomes by Risk Factors, Screening Intervals, and Modalities

No trials reported mortality outcomes for women with specific risk factors besides age, and none compared different screening intervals or modalities. Two observational studies indicated no differences in breast cancer mortality after changing from annual to biennial screening (66) or between annual and triennial screening (67) (Appendix Table 2).

### Advanced Breast Cancer Outcomes

#### Screening Trials

Trials measured cancer severity in terms of clinical stage (0 to IV) (23, 24, 68, 69), number of involved lymph nodes (0, 1 to 3, or ≥4) (18, 20, 25, 26, 70), and tumor size (mm) (19, 21, 26), and these varied across trials. Although “advanced” breast cancer is classified as metastatic disease (stage IV) by the American Joint Committee on Cancer TNM system based on tumor size, lymph node involvement, and presence of metastasis (71), most trials defined advanced breast cancer at much lower thresholds (72).

To combine results, the meta-analysis included the most severe disease categories available from the trials, recognizing that these definitions do not represent equivalent disease stages (Appendix Table 3, available at www.annals.org). These include stage III and IV disease (i.e., regional and metastatic, respectively), size 40 to 50 mm or greater, or 4 or more positive lymph nodes. Combining results on the basis of these definitions indicated no difference with screening for women aged 39 to 49 years (RR, 0.98 [CI, 0.74 to 1.37]) (4 trials [19, 26, 68, 70]) but reduced risk for those aged 50 years or older (RR, 0.62 [CI, 0.46 to 0.83]) (3 trials [21, 26, 68]) (Figure 2).

#### Observational Studies

Five case-series studies compared breast cancer diagnoses in populations of women who had previous screening versus none (73–77). However most studies used thresholds indicating early stages of disease (74, 77) or reported proportions rather than incidence rates (73–77), providing inadequate data to determine the effectiveness of screening (Appendix Table 4, available at www.annals.org).
Treatment-Related Morbidity Outcomes

Screening Trials

A Cochrane review compared treatments between randomized groups in 5 screening trials: the CNBSS-1 and CNBSS-2, the MMST, and the Kopparberg and Stockholm trials (78). In this analysis, women randomly assigned to screening were more likely to have surgical therapy, analyzed as mastectomies and lumpectomies combined (RR, 1.35 [CI, 1.26 to 1.44]) (5 trials) or mastectomies alone (RR, 1.20 [CI, 1.11 to 1.30]) (5 trials) (78). These women were also more likely to have radiation therapy (RR, 1.32 [CI, 1.16 to 1.50]) (2 trials), and less likely to have hormone therapy (RR, 0.73 [CI, 0.55 to 0.96]) (2 trials). Use of chemotherapy was similar between groups (RR, 0.96 [CI, 0.78 to 1.19]) (2 trials) (78).

Observational Studies

Four case-series studies compared breast cancer treatments in populations of women who had previous screening versus none (73–76) (Appendix Table 5, available at www.annals.org). Although studies indicated less extensive surgery, such as fewer total mastectomies and more breast conservation therapies (73–76), and less chemotherapy (73, 74, 76) among screened women, the diagnosis of ductal carcinoma in situ was included in some studies (74, 76), resulting in less intensive therapies overall in screened women.

Advanced Breast Cancer and Treatment-Related Morbidity Outcomes by Risk Factors, Screening Intervals, and Modalities

Five observational studies based on populations in the U.S. Breast Cancer Surveillance Consortium compared breast cancer diagnoses by screening intervals (79–83) (Appendix Table 4). Some analyses indicated no differences between annual and biennial screening in detecting advanced stage disease (79, 80, 83), whereas 2 analyses indicated earlier stages of disease among women aged 40 to 49 years who were screened annually versus biennially (81, 82); this latter finding was confined to women with extreme breast density in one study (82). A randomized trial of annual versus triennial screening indicated detection of more tumors larger than 20 mm in size with triennial screening; however, this threshold indicates early rather than advanced disease (84). Two observational studies of women receiving mammography versus mammography and tomosynthesis indicated no differences in cancer size (85) or node status (85, 86) (Appendix Table 6, available at www.annals.org).

Discussion

A summary of evidence is provided in Table 2. The effectiveness of mammography screening in reducing breast cancer mortality was evaluated by RCTs and observational studies providing fair-quality evidence. Our meta-analysis of 8 randomized trials indicates that breast cancer mortality is generally reduced with screening; however, estimates are not statistically significant for women aged 39 to 49 years and those aged 70 to 74 years, the magnitudes of effect are small, and results differ depending on how cases were accrued in trials. These results differ from our previous estimate (2) because they include updated data from the CNBSS-1 and CNBSS-2, the Swedish Two-County Trial, and the
Age trial and incorporate data by using the longest case accrual methods available from each trial. Observational studies of population-based mammography screening, limited by inherent biases of nonrandomized studies, reported a wide range of reductions in breast cancer mortality. Most studies were conducted in Europe or the United Kingdom and included women aged 50 to 69 years. Meta-analyses indicated a 25% reduction based on 7 incidence-based mortality studies and a 31% reduction based on 7 case–control studies. These results generally concur with our meta-analysis of trials for women aged 50 to 69 years that indicated a statistically significant 22% reduction.

Evidence of breast cancer mortality reduction in observational studies is inconsistent with randomized trials of women in their 40s. Two observational studies indicated 25% to 44% reductions with screening that differ from the non-statistically significant reduction from our meta-analysis of trials. This difference may reflect dissimilarities between participants and nonparticipants of screening programs in nonrandomized studies, as well as assumptions underlying mortality estimates.

All-cause mortality did not differ between randomized groups in meta-analyses of fair-quality trials, regardless of whether trials were analyzed in combined or separate age groups.

Questions about the effectiveness of screening in reducing breast cancer-specific or all-cause mortality on the basis of risk factors, screening intervals, and modalities remain largely unanswered by currently available research. No studies evaluated breast cancer-specific or all-cause mortality outcomes on the basis of risk factors besides age. Although there were no trials of the effectiveness of different screening intervals, 2 observational studies indicated no differences in breast cancer mortality after changing from annual to biennial screening or between annual and triennial screening.

---

**Table 2. Summary of Evidence: Effectiveness of Breast Cancer Screening**

<table>
<thead>
<tr>
<th>Effectiveness of screening in reducing breast cancer-specific and all-cause mortality: differences by age, risk factors, and screening intervals</th>
<th>Studies in Update</th>
<th>Overall Quality</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammography screening reduced breast cancer mortality in RCTs for women aged 39–49 y (RR, 0.85 [95% CrI, 0.75–0.96]; 8 trials), those aged 50–59 y (RR, 0.86 [CrI, 0.75–0.99]; 6 trials), and those aged 60–69 y (0.68 [CrI, 0.54–0.87]; 2 trials); data were limited for women aged 70–74 y</td>
<td>3 RCTs provided updated data in addition to 5 previously published RCTs; 65 observational studies (57 included in 4 systematic reviews, plus 8 additional studies)</td>
<td>Fair</td>
<td>Trials have methodological limitations Observational studies used various methods that introduce potential bias</td>
</tr>
</tbody>
</table>

**Effectiveness of screening in reducing the incidence of advanced breast cancer and treatment-related morbidity: differences by age, risk factors, and screening intervals**

| Not included | 5 RCTs of screening and cancer stage; 1 Cochrane review of 5 RCTs of treatment; 1 RCT of intervals; 14 observational studies | Poor (observational studies) to fair (RCTs) | Definitions of advanced breast cancer were heterogeneous Observational studies were not designed to determine effectiveness |

**Effectiveness of screening in reducing breast cancer-specific and all-cause mortality by screening modality**

| Not included | No studies evaluated this question | NA | NA |

**Effectiveness of screening in reducing the incidence of advanced breast cancer and treatment-related morbidity by screening modality**

| Not included | 2 observational studies | Poor | No RCTs; comparability of groups not known |

BCSC = Breast Cancer Surveillance Consortium; CBE = clinical breast examination; CrI = credible interval; NA = not applicable; RCT = randomized, controlled trial; RR = relative risk; USPSTF = U.S. Preventive Services Task Force.
The effectiveness of breast cancer screening in reducing advanced breast cancer outcomes is supported by less evidence than for mortality. Studies provided heterogeneous measures of breast cancer severity and generally reported early-stage disease. A meta-analysis of trials indicated a statistically significant reduction in advanced disease for women aged 50 years or older who were randomly assigned to undergo screening, but not for women aged 39 to 49 years. This reduction in advanced cancer aligns with reductions in mortality outcomes for women older than 50 years reported in randomized trials and observational studies, but differs from studies of population trends that show no reductions in advanced breast cancer after the introduction of mass screening (87–89).

In a meta-analysis of 5 trials, women randomly assigned to undergo screening were more likely to have surgical and radiation therapy and less likely to have hormone therapy than controls; use of chemotherapy was similar between groups. This finding would be expected, because screening increases detection of ductal carcinoma in situ and early-stage disease that are currently aggressively treated. Observational studies of the impact of screening on diagnosis and treatment of advanced cancer were inadequately designed to determine effectiveness because they generally provided comparisons between screen-detected and non-screen-detected cases and between proportions of different cancer diagnoses for screened versus unscreened women. Comparisons of incidence rates between screening versus nonscreening populations would provide more appropriate measures.

The effectiveness of screening in reducing advanced breast cancer and treatment morbidity on the basis of risk factors, screening intervals, and modalities was also unanswered by current research. The analysis of outcomes based on screening intervals in the U.S. Breast Cancer Surveillance Consortium is limited by the opportunistic nature of screening in the United States. Women choosing short screening intervals probably differ in important ways from those choosing longer intervals. Consequently, comparisons between out-
comes of these 2 types of women may not provide valid measures of effectiveness. Only 2 observational studies compared imaging modalities and found no differences in cancer size or node status between women receiving mammography alone versus mammography and tomosynthesis.

Our review has limitations. First, we included only English-language articles; this could result in language bias, although we did not identify non-English-language studies that otherwise met inclusion criteria in our searches. Second, we only included studies that were applicable to current practice in the United States in order to improve clinical relevance for the USPSTF, excluding studies and limiting relevance to other populations and settings. Third, studies used heterogeneous definitions for advanced breast cancer that did not consider tumor subtypes, and most trials used imaging technologies and treatments that are now outdated, limiting their applicability. Finally, studies were not available for some key questions, specifically for effectiveness based on risk factors, intervals, or other modalities; and the number, quality, and applicability of studies varied widely.

Additional research on the effectiveness of mammography screening with quality-of-life outcomes, as well as morbidity and mortality outcomes, and using current imaging technology and breast cancer treatments would provide further understanding of the implications of routine screening. Data for specific groups of women, particularly older women, or groups based on racial and ethnic background, access to screening, or existence of comorbidities, for example, could further inform screening practices. New technologies, such as tomosynthesis, are becoming more widely used in the United States without definitive studies of their effects on screening outcomes. Studies on the role of additional imaging modalities in screening are required in order to appropriately incorporate these technologies in the screening process.

In conclusion, breast cancer mortality is generally reduced with mammography screening, although estimates are not statistically significant at all ages and the magnitudes of effect are small. Advanced cancer is reduced with screening for women aged 50 years or older.

From the Pacific Northwest Evidence-based Practice Center, Oregon Health & Science University; Veterans Affairs Portland Health Care System; and Providence Cancer 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 AHRQ. No statement in this report should be construed as an official position of AHRQ or the U.S. Department of Health and Human Services.

Acknowledgment: The authors thank Andrew Hamilton, MLS, MS, for conducting literature searches and Spencer Dandy, BS, for assisting with manuscript preparation at the Pacific Northwest Evidence-based Practice Center at Oregon Health & Science University; and Alison Conlin, MD, MPH, and Michael Neuman, MD, at the Providence Cancer Center at Providence Health and Services Oregon, and Arpna Naik, MD, at Oregon Health & Science University for providing medical expertise. They also thank Jennifer Croswell, MD, MPH, at the Agency for Healthcare Research and Quality, and U.S. Preventive Services Task Force members Linda Baumann, PhD, RN; Kirsten Bibbins-Domingo, PhD, MD, MAS; Mark Ebell, MD, MS; Jessica Herzstein, MD, MPH; Michael LeFevre, MD, MSPH; and Douglas Owens, MD, MS.

Grant Support: By the Agency for Healthcare Research and Quality (contract 290-2012-00015-I, Task Order 2).

Disclosures: Drs. Nelson, Fu, Cantor, and Humphrey and Ms. Pappas and Ms. Daeges report grants from AHRQ during the conduct of the study. Authors not named here have disclosed no conflicts of interest. Disclosures can also be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M15-0969.

Requests for Single Reprints: Heidi D. Nelson, MD, MPH, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Mail Code: BICC, Portland, OR 97239; e-mail, nelsonh@ohsu.edu.

Current author addresses and author contributions are available at www.annals.org.

References


**Appendix Figure 1.** Analytic framework and key questions.

- **Women aged ≥40 y**: Women aged 40 years or older.
- **Early detection of invasive breast cancer**: Detection of breast cancer before it becomes invasive.
- **Reduced: Advanced disease treatment morbidity**: Reduction in treatment morbidity.
- **Reduced mortality**: Reduction in breast cancer mortality.
- **Harms of screening**: Potential harms associated with screening.
- **Harms of treatment**: Potential harms associated with treatment.

**Key questions:**

- **For women aged ≥40 years older***:
  1. What is the effectiveness of routine mammography screening in reducing breast cancer–specific and all-cause mortality, and how does it differ by age, risk factor, and screening interval?
  2. What is the effectiveness of routine mammography screening in reducing the incidence of advanced breast cancer and treatment-related morbidity‡, and how does it differ by age, risk factor, and screening interval?
  3. How does the effectiveness of routine breast cancer screening in reducing breast cancer–specific and all-cause mortality vary by different screening modality§?
  4. How does the effectiveness of routine breast cancer screening in reducing the incidence of advanced breast cancer and treatment-related morbidity‡ vary by different screening modality§?

**KQ** = key question.

* Excludes women with preexisting breast cancer; clinically significant BRCA mutations, Li–Fraumeni syndrome, Cowden syndrome, hereditary diffuse gastric cancer, or other familial breast cancer syndromes; high-risk lesions (ductal carcinoma in situ, lobular carcinoma in situ, atypical ductal hyperplasia, atypical lobular hyperplasia); or previous large doses of chest radiation (≥20 Gy) before age 30 y.

‡ Risk factors include family history; breast density; race/ethnicity; menopausal status; current use of menopausal hormone therapy or oral contraceptives; prior benign breast biopsy; and, for women aged >50 y, body mass index.

§ Morbidity includes physical adverse effects of treatment, quality-of-life measures, and other measures of impairment.

§§ Screening modalities include mammography (film, digital, tomosynthesis), magnetic resonance imaging, ultrasonography, and clinical breast examination (alone or in combination).
Appendix Figure 2. Summary of evidence search and selection.

Abstracts of potentially relevant articles identified through MEDLINE and Cochrane databases* \((n = 12,070)\)

Excluded abstracts \((n = 10,036)\)

Full-text articles reviewed \((n = 2,034)\)

Full-text articles excluded \((n = 1,951)\)
- Wrong population: 129
- Wrong intervention: 243
- Wrong outcomes: 532
- Wrong study design: 214
- Wrong publication type: 308
- Included in an included systematic review and not directly used: 68
- Review not meeting inclusion criteria: 125
- Studies outside search dates: 63
- No original data; publication or data set with longer follow-up, more complete data, or same data already included: 30

Included for questions about screening harms
- 10 reviews (134 studies)
- 1 meta-analysis (3 RCTs)
- 46 observational studies
- 2 modeling studies

Included studies \((n = 38)\)†

Effectiveness in reducing mortality, by age, risk factors, and intervals
- 3 RCTs with updated results
- 5 RCTs with no updates
- 4 reviews (57 studies)
- 8 observational studies

Effectiveness in reducing advanced cancer, by age, risk factors, and intervals
- 6 RCTs
- 1 reviews (5 RCTs)
- 14 observational studies

Effectiveness in reducing mortality, by modalities
- No studies

Effectiveness in reducing advanced cancer, by modalities
- 2 observational studies

RCT = randomized, controlled trial.
* Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews.
† Publications may have been used for multiple key questions.
**Appendix Table 1. Mammography Screening Trials**

<table>
<thead>
<tr>
<th>Author (Reference)</th>
<th>Year Trial Began</th>
<th>Setting and Population</th>
<th>Screening Group; Control Group, n*</th>
<th>Method of Randomization</th>
<th>Comparison Groups</th>
<th>Screening Characteristics</th>
<th>Study Duration, y</th>
<th>Longest Follow-up, y</th>
<th>USPSTF Quality Rating and Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIP (29, 65, 68, 90)</td>
<td>1963</td>
<td>New York health plan members aged 40-64 y</td>
<td>30 239; 30 765</td>
<td>Age- and family size-stratified pairs of women were individually randomized by drawing from a list</td>
<td>Mammography + CBE vs. UC</td>
<td>12</td>
<td>4</td>
<td>2</td>
<td>46</td>
</tr>
<tr>
<td>CNBSS-1 and CNBSS-2 (15, 19, 21)</td>
<td>1980</td>
<td>Self-selected participants from 15 centers in Canada aged 40-49 y (CNBSS-1) and 50-59 y (CNBSS-2)</td>
<td>CNBSS-1: 25 214; CNBSS-2: 19 711; 19 694</td>
<td>Individual within blocks stratified by center and 5-year age group after CBE</td>
<td>Mammography + CBE vs. UC (all women prescreened with CBE and instructed in BSE); women 50-59 UC involved annual CBE; all age ≥50 offered screening after trial completed</td>
<td>12</td>
<td>4-5</td>
<td>2</td>
<td>85</td>
</tr>
<tr>
<td>Gothenburg (25, 91, 92)</td>
<td>1982</td>
<td>All women aged 39-59 y born between 1923 and 1944 living in Gothenburg, Sweden</td>
<td>21 650; 29 961</td>
<td>Cluster, based on day of birth for 1923-1935 cohort (18%), by individual for 1936-1944 cohort (82%)</td>
<td>Mammography vs. UC; controls offered screening after 5 y; trial completed after approximately 7 y</td>
<td>18</td>
<td>5</td>
<td>1-2</td>
<td>75</td>
</tr>
<tr>
<td>Stockholm (23, 93)</td>
<td>1981</td>
<td>Residents aged 40-64 y from southeast greater Stockholm, Sweden</td>
<td>40 318; 19 943</td>
<td>Individual, by day of month, ratio of screening to control group 2:1</td>
<td>Mammography vs. UC; controls offered screening after 5 y</td>
<td>24-28</td>
<td>2</td>
<td>1</td>
<td>81</td>
</tr>
<tr>
<td>MMST I and MMST II (24, 94, 95)</td>
<td>1976-1978</td>
<td>All women aged 43-69 y born between 1908 and 1945 living in Malmö, Sweden</td>
<td>MMST I: 21 088; MMST II: 9581; 8212</td>
<td>Individual, within birth year</td>
<td>Mammography vs. UC; controls offered screening after year 14</td>
<td>18-24</td>
<td>9</td>
<td>1-2</td>
<td>70</td>
</tr>
<tr>
<td>Swedish Two-County Trial (26, 96, 97)</td>
<td>1977</td>
<td>Women aged 40-70 y from Östergötland and Kopparberg counties in Sweden</td>
<td>77 080; 55 985</td>
<td>Clusters, based on geographic units; blocks designed to be demographically homogeneous</td>
<td>Mammography vs. UC; controls offered screening after year 7</td>
<td>24-33</td>
<td>3</td>
<td>1</td>
<td>84</td>
</tr>
<tr>
<td>Age(22, 27)</td>
<td>1991</td>
<td>Women aged 39-41 y from 23 National Health Service breast screening units in England, Scotland, and Wales</td>
<td>53 884; 106 956</td>
<td>Individual, stratified by general practitioner group with random number generation 1991-1992; 1992 onward, randomization via Health Authority computer system</td>
<td>Mammography vs. UC; all women offered screening at age 50-52 y</td>
<td>12</td>
<td>4-6, varied by center</td>
<td>2</td>
<td>57</td>
</tr>
</tbody>
</table>

**BSE = breast self-examination; CBE = clinical breast examination; CNBSS = Canadian National Breast Screening Study; HIP = Health Insurance Plan of New York; MMST = Malmö Mammographic Screening Trial; UC = usual care; USPSTF = U.S. Preventive Services Task Force.**

* Numbers of participants in screening and control groups vary by publication.
† Generally effective randomization and comparable groups are assembled initially, but some question remains whether some, although not major, differences occurred in follow-up.
‡ Important differential loss to follow-up or overall high loss to follow-up; adherence <80%.
§ Numbers of participants unclear.
¶ New data since prior recommendation.
¶ Did not maintain comparable groups (includes attrition, crossovers, adherence, contamination).
## Appendix Table 2. Observational Studies of Screening and Mortality Not Included in Systematic Reviews

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Design</th>
<th>Setting</th>
<th>Study Years</th>
<th>Participants, n</th>
<th>Age, y</th>
<th>Participation Rate</th>
<th>Comparison</th>
<th>Adjusted for Previous Breast Cancer</th>
<th>Reduction in Breast Cancer Mortality</th>
<th>Reduction in All-Cause Mortality</th>
<th>Quality Rating and Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coldman et al, 2008 (66)</td>
<td>Time-trend</td>
<td>British Columbia, Canada: 4 cohorts based on date and age at first screening</td>
<td>1988-2005</td>
<td>658 151</td>
<td>40-79</td>
<td>70%</td>
<td>Change from annual to biennial in 1997 for age 50-79 y</td>
<td>NR</td>
<td>Breast cancer deaths (MR pre vs. post): 40.49 vs. 0.67 (95% CI, 0.33-1.37); 50-59 y: 1.06 (CI, 0.74-1.46)</td>
<td>NR</td>
<td>NA</td>
</tr>
<tr>
<td>Coldman et al, 2014 (63)</td>
<td>Incidence-based mortality</td>
<td>Canadian Screening Programs</td>
<td>1990-2009</td>
<td>2 796 472</td>
<td>40-79</td>
<td>85% of Canadians</td>
<td>Women participating in screening vs. not participating</td>
<td>NR</td>
<td>Breast cancer deaths (SMR): 40.49 vs. 0.67 (CI, 0.45-0.67); 50-59 y: 0.60 (CI, 0.49-0.70); 60-69 y: 0.58 (CI, 0.50-0.67); 70-79 y: 0.65 (CI, 0.56-0.74)</td>
<td>NR</td>
<td>NA</td>
</tr>
<tr>
<td>Hellquist et al, 2011 (59)</td>
<td>Prospective cohort</td>
<td>Swedish counties in Mammography Screening of Young Women cohort</td>
<td>1986-2005</td>
<td>620 620</td>
<td>40-49</td>
<td>80%-90% Invited vs. not invited to screen</td>
<td>Yes</td>
<td>Breast cancer deaths (person-years), invited vs. not invited: 619 vs. 1205; RR, 0.74 (CI, 0.66-0.83); Adjusted for attendance: 523 vs. 1,205; RR, 0.71 (CI, 0.62-0.80); NNS during a 10-y period to save 1 life: 1252 (CI, 958-1915)</td>
<td>NR</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Hofvind et al, 2013 (61)</td>
<td>Prospective cohort</td>
<td>Norwegian Breast Cancer Screening Program</td>
<td>1996-2010</td>
<td>699 628</td>
<td>50-69</td>
<td>84% Screened vs. not screened</td>
<td>Unclear</td>
<td>Breast cancer deaths (women years), nonscreened vs. screened Number of deaths: 392/2035 vs. 998/13 162 Adjusted breast cancer mortality: 1.00 vs. 0.39 (CI, 0.35-0.44); Adjusted for self-selection bias: 1.00 vs. 0.57 (CI, 0.51-0.64)</td>
<td>NR</td>
<td>Fair*</td>
<td></td>
</tr>
<tr>
<td>Mook et al, 2011 (62)</td>
<td>Retrospective cohort</td>
<td>The Netherlands</td>
<td>1990-2000</td>
<td>2592</td>
<td>50-69</td>
<td>70%-80% Screened vs. not screened</td>
<td>Yes</td>
<td>Breast cancer mortality (HR), screen-detected vs. not detected: Univariate HR, 0.43 (CI, 0.34-0.53); P &lt; 0.001 Multivariate HR, 0.66 (CI, 0.50-0.86); P = 0.002 Absolute reduction in breast cancer mortality at 10 years of follow-up: 7%</td>
<td>All-cause mortality (HR) Univariate HR: 0.60 (CI, 0.51-0.69); P &lt; 0.001 Multivariate HR: 0.77 (CI, 0.64-0.92); P = 0.005</td>
<td>Poor*</td>
<td></td>
</tr>
<tr>
<td>Parvinen et al, 2011 (67)</td>
<td>Retrospective cohort</td>
<td>Finland, national screening program registry data</td>
<td>1987-2003</td>
<td>14 765</td>
<td>40-49</td>
<td>85% Annual vs. triennial screening</td>
<td>No</td>
<td>Breast cancer mortality (per 100 000 person-years) Triennial: 17.9, RR (reference) Annual: 20.2, RR: 1.14 (CI, 0.99-1.27)</td>
<td>All-cause mortality (per 100 000 person-years) Triennial: 192.6, RR (reference) Annual: 230.9, RR: 1.20 (CI, 0.99-1.46)</td>
<td>Fair†</td>
<td></td>
</tr>
<tr>
<td>Schönberg et al, 2009 (64)</td>
<td>Retrospective cohort</td>
<td>Medical record review at community health centers in the United States</td>
<td>1994-2004</td>
<td>2011</td>
<td>&gt;80</td>
<td>NA</td>
<td>Screened vs. not screened</td>
<td>Yes</td>
<td>Breast cancer deaths: 1 vs. 2 All-cause deaths: 12 vs. 12</td>
<td>All-cause deaths: 12 vs. 12</td>
<td>Fair†</td>
</tr>
</tbody>
</table>

HR = hazard ratio; MR = mortality ratio; NA = not applicable (quality rating criteria unavailable for this study design); NNS = number needed to screen; NR = not reported; RR = relative risk; SMR = standardized mortality ratio.

* Did not maintain comparable groups (includes attrition, crossovers, adherence, contamination).
† Statistical limitations, including low power to detect differences.
### Appendix Table 3. Advanced Breast Cancer Outcomes Reported in Screening Trials

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Stage</th>
<th>Positive Lymph Nodes, n*</th>
<th>Size, mm†</th>
<th>Definition of Advanced Cancer‡</th>
<th>RR for Advanced Cancer (95% CI)§</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIP, 1988 (68)</td>
<td>I, II, III, IV</td>
<td>NR</td>
<td>NR</td>
<td>Stage III–IV</td>
<td>40–49 y: 0.87 (0.48–1.58) 50–64 y: 0.52 (0.31–0.88)</td>
</tr>
<tr>
<td>CNBSS-1, 1992 (18) and</td>
<td>NR</td>
<td>0, 1–3, ≥4</td>
<td>1–9, 10–14, 15–19, 20–39, ≥40</td>
<td>Size ≥40 mm; ≥4 lymph nodes</td>
<td>40–49 y: 1.18 (0.67–2.03) 40–49 y: 2.00 (1.20–3.34)</td>
</tr>
<tr>
<td>CNBSS-2, 1992 (20) and</td>
<td>NR</td>
<td>0, 1–3, ≥4</td>
<td>1–9, 10–14, 15–19, 20–39, ≥40</td>
<td>Size ≥40 mm; ≥4 lymph nodes</td>
<td>50–59 y: 0.75 (0.38–1.46) 50–59 y: 0.91 (0.55–1.49)</td>
</tr>
<tr>
<td>2002 (21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stockholm, 1997 (23)</td>
<td>0, I, II, III–IV</td>
<td>NR</td>
<td>NR</td>
<td>Stage III or greater</td>
<td>40–64 y: 1.15 (0.59–2.07)</td>
</tr>
<tr>
<td>MMST, 1988 (24)</td>
<td>0, I, II, III–IV, III–IV</td>
<td>NR</td>
<td>NR</td>
<td>Stage III or greater</td>
<td>45–70 y: 0.82 (0.56–1.20)</td>
</tr>
<tr>
<td>Swedish Two-County Trial, 1995 (26) and 2003 (69)</td>
<td>I, II, III–IV</td>
<td>0, ≥1</td>
<td>1–9, 10–14, 15–19, 20–29, 30–49, ≥50</td>
<td>Size ≥50 mm</td>
<td>40–49 y: 1.57 (0.63–3.94) 50–74 y: 0.63 (0.45–0.82)</td>
</tr>
<tr>
<td>Age, 2005 (70)</td>
<td>NR</td>
<td>0, 1–3, ≥4</td>
<td>1–9, 10–14, 15–19, 20–29, 30–49, ≥50</td>
<td>Size ≥50 mm; ≥4 lymph nodes</td>
<td>38–49 y: 0.85 (0.57–1.23) 39–49 y: 0.77 (0.53–1.13)</td>
</tr>
</tbody>
</table>

CNBSS = Canadian National Breast Screening Study; HIP = Health Insurance Plan of Greater New York; MMST = Malmö Mammographic Screening Trial; NR = not reported; RR = relative risk.

* Lymph nodes with micrometastases are classified as Stage IB, otherwise ≥1 positive lymph node is classified as stage IIA or higher.

† Size ≥20 mm is classified as stage IIA or higher; size ≥50 mm is classified as stage IIB or higher.

‡ Represents the highest category of disease reported by the trials.

§ Screening vs. control. Only trials reporting results by age (<50 y; ≥50 y) were included in the meta-analysis.
<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Design</th>
<th>Setting</th>
<th>Study Years</th>
<th>Participants, n</th>
<th>Age, y</th>
<th>Comparison</th>
<th>Outcome Measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bussman et al, 2003 (73)</td>
<td>Case series</td>
<td>United States: Kaiser Permanente</td>
<td>1994–2000</td>
<td>247</td>
<td>42–49</td>
<td>Screened vs. not screened</td>
<td>Stage II–IV; III or IV Stage II–IV: 39% (41/105) vs. 52% (74/142); P = 0.06 Stage III or IV: 4% (n = NR) vs. 9% (n = NR); P = NR</td>
<td></td>
</tr>
<tr>
<td>Dittus et al, 2013 (79)</td>
<td>Case series</td>
<td>United States: BCSC data, multisite</td>
<td>1996–2008</td>
<td>4432</td>
<td>40–74</td>
<td>1-y vs. 2-y screening intervals</td>
<td>Stage, size &gt;20 mm; node-positive 2-year vs. 1-year interval No statistically significant differences for stage, size, lymph node positive by weight status</td>
<td></td>
</tr>
<tr>
<td>García Fernández et al, 2014 (74)</td>
<td>Case series</td>
<td>Spain: breast cancer program and regular public health system</td>
<td>2002–2012</td>
<td>904</td>
<td>50–69</td>
<td>Screened vs. not screened</td>
<td>Node-positive; ≥3 nodes positive; size &gt;20 mm Cancer detection rate: 3.8/1000 (475/123,445) vs. 9.4/1000 (382/40,797) Invasive: 80% (419/523) vs. 92% (373/403); P &lt; 0.001 Lymph node positive: 75% (312/419) vs. 57% (204/373); P &lt; 0.001 ≥3 nodes positive: 28% (28/103) vs. 42% (66.156); P &lt; 0.001 Tumor size &gt;20 mm: 16.5% (69/419) vs. 48% (181/223); P &lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Goel et al, 2007 (80)</td>
<td>Case series</td>
<td>United States: Vermont Breast Cancer Surveillance System</td>
<td>1994–2002</td>
<td>1944</td>
<td>&gt;40</td>
<td>1-y vs. 2-y screening intervals</td>
<td>Advanced: either stage IIB or greater; size &gt;20 mm; &gt;1 positive node Advanced: 21% vs. 24%; P = 0.262 No statistically significant differences by age</td>
<td></td>
</tr>
<tr>
<td>Hubbard et al, 2011 (81)</td>
<td>Case series</td>
<td>United States: BCSC data, multisite</td>
<td>1996–2006</td>
<td>4492</td>
<td>40–59</td>
<td>1-y vs. 2-y screening intervals</td>
<td>Stage IIB or greater Adjusted proportion of cancer stage for 2-year vs 1-year intervals Stage III or IV for 40–49 y: 4.8 (95% CI, 1.3–8.4) No statistically significant differences for other stages</td>
<td></td>
</tr>
<tr>
<td>Jensen et al, 2003 (75)</td>
<td>Case series</td>
<td>Denmark and Sweden 1996–1997</td>
<td>2104</td>
<td>50–69</td>
<td>Regions with mammography screening vs. regions without</td>
<td>Stage III or IV; median size Stage III or IV: 8.8% (81/917) vs. 13.6% (162/1187); P &lt; 0.001 Median tumor size (mm): 18 (Malmo) and 17 (Funen) vs. 20 (Aarhus and Northern Jutland); P &lt; 0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kerlikowske et al, 2013 (82)</td>
<td>Case series</td>
<td>United States: BCSC data, multisite</td>
<td>1996–2008</td>
<td>11,474</td>
<td>40–74</td>
<td>1-y vs. 2-y vs. 3-y screening intervals</td>
<td>Stage IIB–IV Adjusted OR for 2-year vs. 1-year intervals Stages IIB–IV in 40-49 y + extreme breast density: 1.89 (95% CI, 1.06–3.41) Tumor size &gt;20 mm in 40-49 y + extreme breast density: 2.39 (95% CI, 1.37–4.18) No statistically significant differences for 50–74 y, 40–49 y without extreme density, or for any comparisons between 3-y vs. 2-y intervals</td>
<td></td>
</tr>
<tr>
<td>Olivotto et al, 1999 (76)</td>
<td>Case series</td>
<td>Canada: Screening Mammography Program of British Columbia</td>
<td>1989–1996</td>
<td>13,636</td>
<td>40–89</td>
<td>Screening attenders vs. nonattenders</td>
<td>Invasive: 88% (17,248/19,446) vs. 92% (75,594/81,946); P &lt; 0.001 Stage III or IV; size &gt;20 mm</td>
<td></td>
</tr>
<tr>
<td>Olsson et al, 2009 (77)</td>
<td>Case series</td>
<td>Sweden: MMST 1961–1991</td>
<td>2478</td>
<td>45–69</td>
<td>Invited to screen vs. not invited</td>
<td>Size &gt;20 mm; node-positive Tumor size &gt;20 mm: 23% vs. 36%; P &lt; 0.05 Lymph node positive: 28% vs. 36%; P &lt; 0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White et al, 2004 (83)</td>
<td>Case series</td>
<td>United States: BCSC data, multisite</td>
<td>1996–2001</td>
<td>7840</td>
<td>40–89</td>
<td>1-y vs. 2-y screening intervals</td>
<td>Stage III or IV, size &gt;20 mm; lymph node-positive Tumor size &gt;20 mm: 22% vs. 24% OR (for 2-y interval vs. 1-y interval) Late stage for invasive cancers only: 0.97 (95% CI, 0.84–1.13) Tumor size &gt;20 mm for invasive: 1.07 (95% CI, 0.92–1.24)</td>
<td></td>
</tr>
</tbody>
</table>

BCSC = Breast Cancer Surveillance Consortium; HIP = Health Insurance Plan of New York; MMST = Malmo Mammographic Screening Trial; NHS = National Health Service; NR = not reported; NS = not significant; OR = odds ratio.
### Appendix Table 5. Observational Studies of Breast Cancer Treatment for Screened and Nonscreened Women

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Design</th>
<th>Setting</th>
<th>Study Years</th>
<th>Participants, n</th>
<th>Age, y</th>
<th>Comparison</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buseman et al, 2003 (73)</td>
<td>Case series</td>
<td>United States: Kaiser Permanente</td>
<td>1994-2000</td>
<td>247</td>
<td>42-49</td>
<td>Screened vs. not screened</td>
<td>Lumpectomy + radiation treatment: 61% (64/105) vs. 57% (81/142); RR, 1.00 (95% CI, 0.75-1.33)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Chemotherapy: 55% (58/105) vs. 61% (86/142); RR, 1.06 (CI, 0.85-1.33)</td>
</tr>
<tr>
<td>García Fernández et al, 2014 (74)</td>
<td>Case series</td>
<td>Spain: breast cancer program and public health system</td>
<td>2002-2012</td>
<td>904</td>
<td>50-69</td>
<td>Screened vs. not screened; includes DCIS (20% vs. 8%; P &lt; 0.001)</td>
<td>Primary treatment: Overall differences, P &lt; 0.001 Conservative surgery: 83% (433/523) vs. 57% (230/403) Radical surgery: 16% (84/523) vs. 41% (163/403) Chemotherapy: 0.4% (2/510) vs. 0.8% (3/394) Sentinel node biopsy: 73% (384/523) vs. 50% (200/403); P &lt; 0.001 Adjuvant treatment: Overall differences, P &lt; 0.001 Chemotherapy: 41% (211/510) vs. 72% (284/394) Hormone therapy: 86% (439/510) vs. 80% (317/394) Radiotherapy: 87% (444/510) vs. 75% (296/394)</td>
</tr>
<tr>
<td>Jensen et al, 2003 (75)</td>
<td>Case series</td>
<td>Denmark and Sweden</td>
<td>1996-1997</td>
<td>2104</td>
<td>50-69</td>
<td>Regions with mammography screening vs. regions without</td>
<td>Overall differences, P &lt; 0.001 Mastectomy: 61% (556/917) vs. 85% (893/1051) Lumpectomy: 32% (295/917) vs. 6.8% (72/1051) Biopsy only: 6.4% (59/917) vs. 8% (84/1051)</td>
</tr>
<tr>
<td>Olivetto et al, 1999 (76)</td>
<td>Case series</td>
<td>Canada: Screening Mammography Program of British Columbia</td>
<td>1989-1996</td>
<td>13 636</td>
<td>40-89</td>
<td>Attendees vs. nonattendees; includes DCIS (12% vs. 8%; P &lt; 0.001)</td>
<td>Definitive breast surgery: Overall differences, P &lt; 0.001 Total mastectomy: 35% (603/1712) vs. 46% (3452/7523) Breast conservation: 65% (1109/1712) vs. 54% (4071/7523) Adjuvant systemic therapy: Overall differences, P &lt; 0.001 Tamoxifen alone: 29% (493/1712) vs. 36% (2694/7523) Chemotherapy: 23% (392/1712) vs. 27% (2060/7523), P &lt; 0.001</td>
</tr>
</tbody>
</table>

DCIS = ductal carcinoma in situ; OR = odds ratio; RR = relative risk.
### Appendix Table 6. Observational Studies of Advanced Cancer Outcomes With Mammography Plus Tomosynthesis

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Design</th>
<th>Setting</th>
<th>Study Years</th>
<th>Participants, n</th>
<th>Age, y</th>
<th>Comparison</th>
<th>Outcome Measures/Definitions</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rose et al, 2013 (86)</td>
<td>Case series</td>
<td>United States: multisite community-based breast center</td>
<td>2011–2012</td>
<td>DM: 18,202 DM + T: 10,878</td>
<td>&gt;18</td>
<td>DM vs. DM + T</td>
<td>Cancer detection rate; positive nodes</td>
<td>Cancer detection rate: 4.0 vs. 5.4/1000; NS Positive nodes: 4 vs. 6; ( P = 0.84 )</td>
</tr>
<tr>
<td>Skaane et al, 2013 (85)</td>
<td>Postintervention series</td>
<td>Norway: Oslo screening program;</td>
<td>2010–2011</td>
<td>12,631</td>
<td>50–69</td>
<td>DM vs. DM + T (biennial screening)</td>
<td>Cancer detection rate; positive nodes; size ≥20 mm</td>
<td>Cancer detection rate: 6.1/1000 vs. 8.0/1000; ( P = 0.001 ) Positive nodes: 9 vs. 13; NS Size ≥20 mm: 12 vs. 15; NS</td>
</tr>
</tbody>
</table>

DM = digital mammography; NS = not statistically significant; T = tomosynthesis.