Annals of Internal Medicine

REVIEW

Harms of Breast Cancer Screening: Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation

Heidi D. Nelson, MD, MPH; Miranda Pappas, MA; Amy Cantor, MD, MPH; Jessica Griffin, MS; Monica Daeges, BA; and Linda Humphrey, MD, MPH

Background: In 2009, the U.S. Preventive Services Task Force recommended biennial mammography screening for women aged 50 to 74 years and selective screening for those aged 40 to 49 years.

Purpose: To review studies of screening in average-risk women with mammography, magnetic resonance imaging, or ultrasonography that reported on false-positive results, overdiagnosis, anxiety, pain, and radiation exposure.

Data Sources: MEDLINE and Cochrane databases through December 2014.

Study Selection: English-language systematic reviews, randomized trials, and observational studies of screening.

Data Extraction: Investigators extracted and confirmed data from studies and dual-rated study quality. Discrepancies were resolved through consensus.

Data Synthesis: Based on 2 studies of U.S. data, 10-year cumulative rates of false-positive mammography results and biopsies were higher with annual than biennial screening (61% vs. 42% and 7% vs. 5%, respectively) and for women aged 40 to 49 years, those with dense breasts, and those using combination hormone therapy. Twenty-nine studies using different methods reported overdiagnosis rates of 0% to 54%; rates from randomized trials were 11% to 22%. Women with false-positive results reported more anxiety, distress, and breast cancer–specific worry, although results varied across 80 observational studies. Thirty-nine observational studies indicated that some women reported pain during mammography (1% to 77%); of these, 11% to 46% declined future screening. Models estimated 2 to 11 screening-related deaths from radiation-induced cancer per 100 000 women using digital mammography, depending on age and screening interval. Five observational studies of tomosynthesis and mammography indicated increased biopsies but reduced recalls compared with mammography alone.

Limitations: Studies of overdiagnosis were highly heterogeneous, and estimates varied depending on the analytic approach. Studies of anxiety and pain used different outcome measures. Radiation exposure was based on models.

Conclusion: False-positive results are common and are higher for annual screening, younger women, and women with dense breasts. Although overdiagnosis, anxiety, pain, and radiation exposure may cause harm, their effects on individual women are difficult to estimate and vary widely.

Primary Funding Source: Agency for Healthcare Research and Quality.

Ann Intern Med. 2016;164:XXX-XXX. doi:10.7326/M15-0970 **www.annals.org**

For author affiliations, see end of text.

This article was published at www.annals.org on 12 January 2016.

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, 3). 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 those aged 75 years or older (1).

Although there is general consensus that mammography screening is beneficial for many women, benefits must be weighed against potential harms to determine the net effect of screening on individual women. Determining the balance between benefits and harms is complicated by several important considerations that are unresolved, including defining and quantifying potential harms; the optimal ages at which to begin and end routine screening; the optimal screening intervals; appropriate use of various imaging modalities, including supplemental technologies; values and preferences of women in regards to screening; and how all of these considerations vary depending on a woman's risk for breast cancer.

This systematic review updates evidence for the USPSTF on the harms of breast cancer screening, including false-positive mammography results, overdiagnosis, anxiety, pain during procedures, and radiation exposure, and how these adverse effects vary by age, risk factor, screening interval, and screening modality. Systematic reviews of the effectiveness of screening (4), performance characteristics of screening methods (5), and the accuracy of breast density determination and use of supplemental screening technologies (6) are provided in additional reports.

METHODS

Scope, Key Questions, and Analytic Framework

The USPSTF determined the scope and key questions for this review by using established methods (7, 8). A standard protocol was developed and publicly posted on the USPSTF Web site. A technical report further describes the methods and includes search strategies and additional information (4).

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 harms of routine breast cancer screening and how they

See also:

Related articles 1
Editorial comment 2

differ by age, risk factor, screening interval, and screening modality (mammography [film, digital, or tomosynthesis], magnetic resonance imaging [MRI], and ultrasonography). Harms include false-positive and false-negative mammography results, overdiagnosis, anxiety and other psychological responses, pain during procedures, and radiation exposure. Overdiagnosis refers to women receiving a diagnosis of ductal carcinoma in situ (DCIS) or invasive breast cancer when they have abnormal lesions that are unlikely to become clinically evident during their lifetime in the absence of screening. Overdiagnosed women may be harmed by unnecessary procedures and treatments as well as by the burden of receiving a cancer diagnosis.

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 recommendations for preventive services (preexisting breast cancer or high-risk breast lesions, hereditary genetic syndromes associated with breast cancer, and 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 (9) (described in Appendix Figure 1).

Data Sources and Searches

A research librarian conducted electronic searches of the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, and Ovid MEDLINE through December 2014 for relevant studies and systematic reviews. Searches were supplemented by references identified from additional sources, including reference lists and experts. Studies of harms included in the previous systematic review for the USPSTF (2, 3) were also included.

Study Selection

Two investigators independently evaluated each study to determine eligibility based on prespecified inclusion criteria. Discrepancies were resolved through consensus.

We included recently published systematic reviews; randomized, controlled trials (RCTs); and observational studies of prespecified harms. When available, studies providing outcomes specific to age, risk factors, screening intervals, and screening modalities were preferred over studies providing general outcomes. Studies that were most clinically relevant to practice in the United States were selected; 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 with designs ranked higher in the study design–based hierarchy of evidence were emphasized because they are less susceptible to bias (for example, RCTs were chosen 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 another. Two investigators independently applied criteria developed by the USPSTF (7, 8) to rate the quality of each RCT, cohort study, case–control study, and systematic review as good, fair, or poor; criteria to rate studies with other designs included in this review are not available. Discrepancies were resolved through consensus.

Data Synthesis

Studies meeting inclusion criteria were qualitatively synthesized. Most studies in this review had designs for which quality rating criteria are not available, which limited data synthesis. When possible, we assessed the aggregate internal validity (quality) of the body of evidence for each key question (good, fair, or poor) by using methods developed by the USPSTF based on the number, quality, and size of studies; consistency of results between studies; and directness of evidence (7, 8).

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 frameworks, 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 12 004 abstracts identified by searches and other sources, 59 studies met inclusion criteria for key questions in this report, including 10 systematic reviews of 134 studies and 49 additional studies (Appendix Figure 2, available at www.annals.org).

False-Positive Mammography Results

Two new observational studies estimated the cumulative probability of false-positive results after 10 years of screening with film and digital mammography, based on data from the Breast Cancer Surveillance Consortium, a large population-based database in the United States (Appendix Table 1, available at www.annals.org) (10, 11). When screening began at age 40 years, the cumulative probability of receiving at least 1 false-positive mammography result after 10 years was 61% (95% CI, 59% to 63%) with annual screening and 42% (CI, 41% to 43%) with biennial screening (10). Es-

timates were similar when screening began at age 50 years. The cumulative probability of receiving a biopsy recommendation due to a false-positive mammography result after 10 years of screening was 7% (CI, 6% to 8%) with annual screening versus 5% (CI, 4% to 5%) with biennial screening for women who initiated screening at age 40 years and 9% (CI, 7% to 12%) with annual screening versus 6% (CI, 6% to 7%) with biennial screening for those who began at age 50 years.

In a separate analysis, rates of false-positive mammography results were highest among women receiving annual mammography who had extremely dense breasts and either were aged 40 to 49 years (65.5%) or used combination hormone therapy (65.8%) (11). The highest rates of biopsy due to false-positive mammography results were related to similar characteristics and ranged from 12% to 14%. Rates of false-positive mammography results were lower among women aged 50 to 74 years who were receiving biennial or triennial mammography and had breasts with scattered fibroglandular densities (39.7% and 21.9%, respectively) or almost entirely fat breast density (17.4% and 12.1%, respectively), regardless of estrogen use.

Overdiagnosis

A meta-analysis of 3 RCTs (13, 14), a systematic review of 13 observational studies (15), and 18 new individual studies (16–33) of overdiagnosis were identified for this update (4) (Appendix Table 2, available at www.annals.org). Estimates were primarily based on screening trials, screening programs and registries, or modeled data. Studies differed by patient populations; screening and follow-up times; screening policies, uptake, and intensity; and underlying cancer incidence trends. In addition, at least 7 different measures of overdiagnosis were reported (19). Estimates differed in their numerators and denominators, whether they included both invasive cancer and DCIS, their assumptions about lead time and progression of invasive cancer and DCIS, and whether they reported relative or absolute changes.

Various methods were used to estimate overdiagnosis. The most common methods determined the difference in cancer incidence in the presence and absence of screening (observed excess incidence approach) or made inferences about the lead time or natural history of breast cancer and estimated the corresponding frequency of overdiagnosis (lead-time approach) (35). How differences in study characteristics, measures, and methods affect estimates of overdiagnosis has been well-described (13, 14, 19, 35–37), yet there is no consensus about the appropriate approach (14) and there are no quality rating criteria to evaluate studies.

Estimates From RCTs

Data from 3 RCTs that did not screen control participants at the end of the trials were considered to be the least biased estimates of overdiagnosis in a comprehensive review (13, 14). The Malmö I trial and the Canadian National Breast Screening Study (CNBSS-1 and CNBSS-2) provided estimates from randomized comparison groups with follow-up that extended sufficiently beyond the screening period to differentiate earlier diagnosis from overdiagnosis (13). However, their approaches differed: The Malmö I trial included all breast cancer cases, and the Canadian trials included only those detected by screening.

Results of the Malmö I trial (34) and the 2 Canadian trials (38, 39) were used to compare the excess incidence of breast cancer (both invasive cancer and DCIS) in the screening population with the incidence in the absence of screening. Overdiagnosis was estimated at 10.7% (CI, 9.3% to 12.2%) (13, 14) when only cases identified during the screening period were included and 19.0% (CI, 15.2% to 22.7%) when cases identified throughout screening and follow-up were included. Estimates for women aged 40 to 49 years in CNBSS-1 (12.4% for shorter accrual and 22.7% for longer accrual) were higher than for those aged 50 to 59 years in CNBSS-2 (9.7% and 16.0%, respectively) and those aged 55 to 69 years in the Malmö I trial (10.5% and 18.7%, respectively). Recently published long-term follow-up of the 2 Canadian trials (15 years after enrollment) indicated a 22% overdiagnosis rate for invasive cancer for the combined age groups (31).

Estimates From Observational Studies

Unadjusted estimates from 13 observational studies included in a systematic review indicated overdiagnosis rates ranging from 0% to 54%, and 6 studies that adjusted for breast cancer risk and lead time indicated rates ranging from 1% to 10% (15). Estimates from other studies fall within this overall broad range.

Anxiety, Distress, and Other Psychological Responses

Four systematic reviews of 70 unique studies (40–43) (Table 1) and 10 additional observational studies (44–53) (Table 2) published after the systematic reviews described adverse psychological effects of screening. Although several studies met criteria for fair or good quality, most were limited by enrollment of small numbers of narrowly selected participants, use of various self-reported measures, differential attrition or response rates, and low clinical applicability. No studies provided results by age, risk factor, screening interval, or screening modality.

Results of systematic reviews indicated that women who received clear communication of their negative mammography results had minimal anxiety, whereas those recalled for further testing had more anxiety, breast cancer–specific worry, and distress (40, 42, 54–57). Some women had persistent anxiety despite eventual negative results (56, 58–61), whereas some showed only transient anxiety (54, 62–68). Among studies that evaluated reattendance rates, 2 studies reported that women with false-positive results were less likely to return for their next screening mammography

Table 1. Systematic Reviews of Psychological Harms of Breast Cancer Screening

Author, Year (Reference)	Inclusion Criteria	Searches	Studies, *n*	Participants, *n*
New studies				
Bond et al, 2013 (43)	Studies in the United Kingdom comparing women with FP vs. normal screening mammograms	Multiple databases through November 2011	7*	3168 (psychological harms); 151 490 (screening reattendance)
Hafslund and Nortvedt, 2009 (42)	Studies of women aged 40 to 74 y not at high risk invited to mammography screening	Multiple databases; January 1995 to July 2007	17†	18 097
2009 review				
Brett et al, 2005 (40)	Studies of the psychological effect of mammography screening	Multiple databases; 1982 to 2003	54	NR
Brewer et al, 2007 (41)	Studies comparing women with FP vs. normal screening mammograms	Multiple databases through September 2006	23	313 967

FP = false-positive; NR = not reported; RR = risk ratio.
* 5 studies were included in ≥1 of the systematic reviews included in the 2009 review.
† 13 studies were included in ≥1 of the systematic reviews included in the 2009 review.

(56, 69) and 2 studies reported no differences (70, 71). One study reported an increase in reattendance when women were given letters tailored to their last screening result (risk ratio, 1.10 [CI, 1.00 to 1.21]) (72).

Five new observational studies compared psychological outcomes in women receiving false-positive results versus those receiving normal results (44, 46–48, 50) and reported findings similar to those of the reviews. Women with false-positive results had more breast cancer–specific worry (49% vs. 10%; $P < 0.0001$), more worries that affected mood or daily activities (31% vs. 2%; $P < 0.0001$) (48), and lower mental functioning (mean mental functioning score on the Short Form-36 at 6 months, 80.6 vs. 85.0; $P = 0.03$) and vitality (mean vitality score on the Short Form-36 at 6 months, 70.3 vs. 77.0; $P = 0.02$) (50). They also had increased measures of depression (mean score on the depression subscale of the Hospital Anxiety and Depression Scale at 6 months, 3.2 vs. 2.4; $P = 0.045$); however, scores were below clinical thresholds for depression (50). An analysis of racial subgroups in a large study indicated increased depression scores among nonwhite women with false-positive results (odds ratio, 3.23 [CI, 1.32 to 7.91]) (44). Three studies found lower reattendance rates for women with false-positive results (51, 52) or biopsies (51, 53), but reattendance sometimes varied by specific circumstances, such as age or type of biopsy (51).

Pain During Procedures

Two systematic reviews included 39 unique studies of pain associated with screening procedures (73, 74), and a separate systematic review included 7 trials of interventions to reduce pain (75) (**Appendix Table 3**, available at www.annals.org). Results indicated that many women had pain (range, 1% to 77%) but few considered it a deterrent to future screening (73). In these studies, pain was associated with stage of the menstrual cycle, anxiety, and the anticipation of pain.

In a review of studies of pain or discomfort after screening mammography and their effect on screening reattendance (74), actual nonreattendance due to concerns about pain ranged from 11% to 46% (5 studies) and intended future nonreattendance ranged from 3% to 18% (2 studies). Fifteen studies that did not directly ask about reasons for nonreattendance found no differences in actual reattendance between women who had pain and those who did not (risk ratio, 1.38 [CI, 0.94 to 2.02]) (5 studies) (74). However, nonreattenders had significantly higher pain scores than reattenders in 2 of 3 studies (76–78). Two studies reported lower intent to reattend among women with pain, whereas 3 others

Table 1–Continued

Outcomes in Women With FP vs. Normal Results				Quality Rating	Limitations
Screening Reattendance	**Anxiety**	**Depression**	**Breast Cancer–Specific Worry/Distress**		
Lower with FP result (2 studies) No difference (2 studies) Higher with FP result if given tailored letters (1 study)	No difference (2 studies)	No difference (2 studies)	Higher with FP result (3 studies)	Good	Unclear whether the quality of studies was considered in the formulation of conclusions
NR	Higher with FP result (15 studies)	NR	Higher with FP result (15 studies)	Fair	Unclear whether the quality of studies was considered in the formulation of conclusions; did not report whether studies were dual-reviewed and dual-abstracted; conflicts of interest were not reported
NR	Higher with FP result (14 studies)	NR	Higher with FP result (9 studies)	Fair	Conflicts of interest and quality rating of studies were not reported
United States: lower with FP result (RR, 1.07 [95% CI, 1.02 to 1.12]) (5 studies) Canada: lower with normal result (RR, 0.63 [CI, 0.50 to 0.80]) (2 studies) Europe: no differences (RR, 0.97 [CI, 0.93 to 1.01]) (5 studies)	Higher with FP result (4 studies) No differences (4 studies) Conflicting results over time (3 studies)	Lower with FP result (1 study) No differences (7 studies) Conflicting results based on measure (1 study)	Higher with FP result (4 studies) No differences (3 studies) Conflicting results over time (2 studies)	Fair	Conflicts of interest were not reported; did not formally assess study quality with prespecified criteria

reported no differences in intended reattendance and pain (79–83).

A systematic review of trials of interventions to reduce pain associated with mammography screening (75) found that providing verbal or written information to women reduced discomfort in 2 studies (84, 85) but not in a third (86). Studies of different breast compression strategies (87, 88) or premedication with acetaminophen (89) indicated no differences in discomfort, whereas use of a breast cushion reduced pain (90).

Radiation Exposure

No studies directly measured the association between radiation exposure from mammography screening and the incidence of breast cancer and death. Two-view digital mammography and screen-film mammography involve average mean glandular radiation doses of 3.7 and 4.7 mGy, respectively, and are considered to provide low-dose, low-energy radiation exposure.

Two modeling studies provided estimates of radiation exposure, breast cancer incidence, and death (91, 92) (**Appendix Table 4**, available at www.annals.org). A model predicting the number of breast cancer cases attributable to the radiation dose of a single typical digital mammogram estimated that the number of deaths due to radiation-induced cancer ranged from 2 per 100 000 in women aged 50 to 59 years screened biennially to 11 per 100 000 in those aged 40 to 59 years screened annually (92).

Differences Between Screening Modalities

Six observational studies compared false-positive recall rates with screening using mammography and tomosynthesis (93–97) or clinical breast examination (98) versus mammography alone (**Appendix Table 5**, available at www.annals.org). No studies evaluated MRI screening in women who were not at high risk for breast cancer.

Four of 5 studies showed statistically significantly lower rates of recall for tomosynthesis and mammography than for mammography alone (93–97). Although recalls were reduced by 16 per 1000 women (CI, −18 to −14 recalls; $P < 0.001$) in one U.S. study, biopsies increased by 1.3 per 1000 women (CI, 0.4 to 2.1 biopsies; $P = 0.004$) (93). A smaller U.S. study showed reduced recall rates with tomosynthesis and mammography versus mammography alone after controlling for age, breast density, and breast cancer risk (adjusted odds ratio, 0.62 [CI, 0.55 to 0.70]; $P < 0.0001$) (97), whereas another study indicated no reductions (94). Two European studies also reported lower rates of recall for women screened with tomosynthesis and mammography (1% vs. 2% [$P < 0.0001$] [95] and 53 vs. 61 per 1000 women [$P = 0.001$] [96]).

Table 2. Results of New Studies of Psychological Harms of Breast Cancer Screening

Author, Year (Reference)	Study Design	Population	Comparisons (Number of Participants)	Measures
Schou Bredal et al, 2013 (49)	Before–after	Women recalled in a screening program in Norway	A: At recall (640) B: 4 wk later	HADS (score ≥11)
Brodersen and Siersma, 2013 (46)	Nested case–control	Screening programs in Denmark	A: FP (272) B: Normal (864) C: TP (174)	COS-BC
Espasa et al, 2012 (48)	Case–control	Screening program in Spain	A: FP (100) B: Normal (50)	HADS, structured interview
Fitzpatrick et al, 2011 (51)	Retrospective cohort	Screening program in the United Kingdom	A: FP (9746) B: Normal (148 589)	Reattendance
Gibson et al, 2009 (44)	Prospective cohort	New Hampshire Mammography Network and the NHWH study	A: FP (2107) B: Normal (11 384)	WHQ
Hafslund et al, 2012 (50)	Nested case–control	Screening programs in Norway	A: FP (128) B: Normal (195)	SF-36, HADS
Keyzer-Dekker et al, 2012 (45)	Prospective cohort	Women with abnormal results in the Netherlands	A: First screen recalls (186) B: Repeated screen recalls (296)	STAI, NEO-FFI, CES-D, WHOQOL
Klompenhouwer et al, 2014 (52)	Retrospective cohort	Screening program in the Netherlands	A: Normal screen (373 474) B: First screen recalls (6672) C: Repeated screen recalls for different lesion (161) D: Repeated screen recalls for same lesion (89)	Reattendance
Maxwell et al, 2013 (53)	Retrospective cohort	Screening program in the United Kingdom	First screening: A: Open biopsy (110) B: Needle sampling (1374) C: No tissue sampling (2703) Repeated screening: A: Open biopsy (199) B: Needle sampling (1052) C: No tissue sampling (4009)	Reattendance – –
Tosteson et al, 2014 (47)	Nested case–control	Women participating in the DMIST in the United States	A: FP (494) immediate B: FP 1 y after C: Normal (534) immediate D: Normal 1 y after	STAI, EuroQol EQ-5D – –

CES-D = Center for Epidemiologic Studies Depression Scale; COS-BC = Consequences of Screening in Breast Cancer; DMIST = Digital Mammographic Imaging Screening Trial; FP = false-positive; HADS = Hospital Anxiety and Depression Scale; NA = not applicable; NEO-FFI = Neuroticism-Extraversion-Openness Five-Factor Inventory; NHWH = New Hampshire Women for Health; NR = not reported; QOL = quality of life; SF-36 = Short Form-36; STAI = State-Trait Anxiety Inventory; TP = true-positive; WHOQOL = World Health Organization Quality of Life; WHQ = Women's Health Questionnaire.
* Both groups improved over time.

Women receiving mammography and clinical breast examination had more recalls than those receiving mammography alone in a study from Canada (8.7% vs. 6.5%; 55 additional recalls per 10 000 women) (98).

DISCUSSION

A summary of the evidence is provided in Table 3. Two large observational studies of women screened in the Breast Cancer Surveillance Consortium provided good-quality evidence about cumulative rates of false-positive mammography results and biopsies over 10 years. In these studies, rates were higher with annual than biennial screening (mammography, 61% vs. 42%; biopsy, 7% vs. 5%) and for women with heterogeneously or extremely dense breasts, those aged 40 to 49 years, and those using combination hormone therapy. These results are consistent with those of an earlier study indicating cumulative 10-year rates of false-positive mammography results of 49% overall and 56% for women aged 40 to 49 years, with an overall biopsy rate of 19% (12). The results of these highly clinically applicable studies can be used to inform women of the likelihood of false-positive results and additional procedures with mammography screening in the United

Table 2–Continued

Outcomes					Quality Rating	Limitations
Screening Reattendance	**Anxiety**	**Depression**	**Breast Cancer–Specific Worry**	**General QOL**		
NR	No difference	No difference	NR	NR	NA	Study design not amenable to quality rating
NR	Immediate: higher for A + C vs. B; no difference for A vs. C 3 y after: higher for C vs. A + B and A vs. B	NR	No difference	NR	Good	FP group significantly younger ($P < 0.05$)
NR	No difference	No difference	Higher for FP vs. normal	NR	Fair	Enrolled selected group of women; did not control for confounders
Decreased: women aged >55 y, open biopsy, longer time to diagnosis Increased: repeated screens, screened in mobile unit	NR	NR	NR	NR	Fair	Did not control for confounders; unclear how women were selected; baseline data not provided for groups of interest
NR	NR	Higher for nonwhite with FP vs. normal	NR	NR	Fair	Unclear how women were selected; baseline data not provided for groups of interest; outcomes self-reported
NR	No difference	More cases for FP vs. normal	NR	Lower for FP vs. normal	Fair	Enrolled selected group of women; higher response rate in control group
NR	No difference*	No difference*	NR	NR	Fair	Outcomes self-reported; older women in repeated screen group; did not report attrition
A: 93.2% B: 65.4% C: 56.7% D: 44.3% All recalled groups combined: 44.3%	NR	NR	NR	NR	Fair	Did not control for confounders; baseline data not provided for groups of interest
Increased for C but no change for A or B	NR	NR	NR	NR	Fair	Did not control for confounders; baseline data not provided for groups of interest
Decreased for A and B but no change for C	NR	NR	NR	NR	–	–
NR	Decreased from A to B	NR	NR	No difference	Good	FP group significantly younger ($P < 0.05$)
NR	No difference	NR	NR	No difference	–	–

States, particularly for women with characteristics associated with the highest rates of false-positive results.

Despite much research, the evidence for determining overdiagnosis is poor. There is no consensus definition of overdiagnosis, and there are no criteria on which to base critical appraisal of studies. Studies are highly heterogeneous, and estimates vary depending on the analytic approach. Possibly the least biased estimates were derived from 3 RCTs that indicated rates of 11% to 22%. Unadjusted estimates from 13 observational studies ranged from 0% to 54%, and 6 studies that adjusted for breast cancer risk and lead time found rates ranging from 1% to 10%. Until methodological standards for estimating overdiagnosis are more clearly defined, the correct estimate is uncertain.

Although overdiagnosis is an important outcome of screening, it is difficult to evaluate in individual women because it is based on knowing whether a specific lesion will progress and what its effect will be on a woman's health. Women who are overdiagnosed can be harmed by unnecessary procedures and treatments and by the burden of receiving a cancer diagnosis. The introduction of technology capable of detecting even smaller suspicious lesions may also lead to increased overdiagnosis. Understanding the concept of overdiagnosis is important to appropriately inform women about the benefits and harms of screening despite current limitations in determining its effect on individual women.

The effect of screening on anxiety and pain is supported by fair-level evidence that includes a large number of predominantly descriptive observational studies. In general, women with false-positive results have more

Table 3. Summary of Evidence

Primary Findings From Previous USPSTF Reviews	Number and Type of Studies in Update	Overall Quality	Limitations	Consistency	Applicability	Summary of Findings
False-positive and false-negative results						
Younger women had higher rates of false-positive mammography results per screening cycle. Cumulative 10-y rates for false-positive mammography results were 49% overall and 56% for ages 40–49 y; cumulative 10-y rate of biopsies due to false-positive mammography results was 19% (based on 1 observational study).	2 observational studies of women screened in the United States	Good	Not all risk factors were examined.	Consistent	Good	10-y cumulative rates of false-positive mammography results and biopsies were higher with annual vs. biennial screening (61% vs. 42% and 7% vs. 5%, respectively) and for women with heterogeneously or extremely dense breasts, those aged 40–49 y, and those using combination hormone therapy.
Overdiagnosis						
Estimates of overdiagnosis ranged from 0% to 50% (based on 1 systematic review and 8 studies).	1 meta-analysis of 3 trials; 1 systematic review of 13 studies; 18 individual studies	Poor	No established definition or method to determine overdiagnosis; studies were highly heterogeneous, and estimates varied depending on the analytic approach.	Inconsistent	Poor	Estimates of overdiagnosis ranged from 0% to 54% overall and from 11% to 22% in randomized trials.
Anxiety and distress						
Many women have anxiety with mammography, but it is generally transient and is not a deterrent to future screening (based on 2 systematic reviews of 77 observational studies).	2 systematic reviews of 24 studies; 10 observational studies	Fair	Studies used different outcome measures and thresholds; effects based on age, risk factors, and screening intervals were not determined.	Consistent	Fair	Women with false-positive results had more anxiety, distress, and breast cancer–specific worry than those with negative results, particularly those who had biopsies, fine-needle aspirations, and early recall; distress persisted for some women but was transient for others. Some women with false-positive results did not return for screening, although some studies showed no differences in reattendance.
Pain						
Many women have pain with mammography, but it is generally transient and is not a deterrent to future screening (based on 1 systematic review of 22 observational studies of pain). Pain could be reduced by providing information to patients or using breast cushions (based on 1 systematic review of 7 trials of interventions to reduce pain).	1 systematic review of 20 observational studies of pain	Fair	Studies used different outcome measures and thresholds; effects based on age, risk factors, and screening intervals were not determined.	Consistent	Fair	Although many women had pain during mammography (1% to 77%), the proportion of those experiencing pain who did not attend future screening varied (11% to 46%).
Radiation exposure						
No studies	2 modeling studies of radiation exposure	Poor	No studies directly measured associations between radiation exposure from mammography screening and breast cancer incidence and death.	Consistent	Poor	Models estimated 2 to 11 deaths per 100 000 women due to radiation-induced cancer from screening with digital mammography, depending on age and screening intervals.

Continued on following page

Table 3–Continued

Primary Findings From Previous USPSTF Reviews	Number and Type of Studies in Update	Overall Quality	Limitations	Consistency	Applicability	Summary of Findings
Harms of screening, by modality						
Not included	5 observational studies of tomosynthesis and 1 of clinical breast examination combined with mammography	Poor	No randomized trials; comparability of groups was not reported; biopsy rates and outcomes were not uniformly reported.	Consistent	Fair	A U.S. study found that tomosynthesis plus mammography resulted in a decrease of 16 recalls and an increase of 1.3 biopsies per 1000 women compared with mammography alone. A Canadian study found that mammography plus clinical breast examination resulted in an increase of 55 recalls per 10 000 women compared with mammography alone.

USPSTF = U.S. Preventive Services Task Force.

anxiety and distress than those with normal results. Anxiety lessens over time for most women but persists for others, and some women with false-positive results do not attend subsequent screenings. Although many women have pain during mammography, the proportion of those who do not attend subsequent screenings varies. Studies indicate that the experiences of false-positive results and pain during mammography differ widely among women but are important for many of them. Additional efforts to reduce false-positive results and improve how they are communicated and to recognize and reduce pain during procedures could improve the balance of benefits and harms of screening for many women.

The harms of radiation exposure from mammography screening are based on only 2 modeling studies. The number of deaths due to radiation-induced cancer from screening with digital mammography was estimated to be 2 to 11 per 100 000 women, depending on age and screening intervals. As imaging technologies change, this estimate could improve or worsen depending on the uptake of supplemental imaging with tomosynthesis as well as additional imaging for false-positive results. Reducing radiation exposure through more effective imaging is an important area of future research.

Five observational studies described false-positive results with the use of tomosynthesis. This evidence is limited by the lack of randomized trials, uncertainty about the comparability of comparison groups, and differences in outcome measures. A U.S. study comparing tomosynthesis and mammography versus mammography alone reported a significant reduction of 16 recalls but an increase of 1.3 biopsies per 1000 women. Available studies of screening with MRI or ultrasonography focus on high-risk women and are outside the scope of this systematic review. No randomized trials of the efficacy of the different imaging technologies for breast cancer screening have been published, and evidence on their benefits and harms for screening recommendations is lacking.

Limitations of this review include the use of English-language articles only, which could have resulted in language bias, although we did not identify non–English-language studies that otherwise met inclusion criteria in our searches. We included only studies that are applicable to current practice in the United States to improve clinical relevance for the USPSTF. The number, quality, and applicability of studies varied widely, and most studies were observational, with designs for which quality rating criteria are not available.

In conclusion, false-positive results are common and lead to additional imaging and biopsies, particularly with annual screening and among younger women and those with dense breasts. Although overdiagnosis, anxiety, pain, and radiation exposure may cause harm, their effects on individual women are difficult to estimate and vary widely.

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.

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; Alison Conlin, MD, MPH, and Michael Neuman, MD, at the Providence Cancer Center at Providence Health & Services Oregon, and Arpana Naik, MD, at Oregon Health & Science University for providing medical expertise; Jennifer Croswell, MD, MPH, at the AHRQ; and USPSTF 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.

Financial Support: By the AHRQ (contract 290-2012-00015-I, task order 2), Rockville, Maryland.

Disclosures: Drs. Nelson, Cantor, and Humphrey; Ms. Pappas; Ms. Griffin; and Ms. Daeges report grants from AHRQ during the conduct of this study. Forms can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M15-0970.

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

1. **U.S. Preventive Services Task Force.** Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2009;151:716-26. [PMID: 19920272] doi:10.7326/0003-4819-151-10-200911170-00008

2. **Nelson HD, Tyne K, Naik A, Bougatsos C, Chan B, Nygren P, et al.** Screening for Breast Cancer: Systematic Evidence Review Update for the US Preventive Services Task Force. AHRQ Publication No. 10-05142-EF-1. Rockville, MD: Agency for Healthcare Research and Quality; 2009. [PMID: 20722173] Accessed at www.ncbi.nlm.nih.gov/pubmed/20722173 on 19 September 2014.

3. **Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L; U.S. Preventive Services Task Force.** Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med. 2009;151:727-37. [PMID: 19920273] doi:10.7326/0003-4819-151-10-200911170-00009

4. **Nelson HD, Cantor A, Humphrey L, Fu R, Pappas M, Daeges M, et al.** Screening for Breast Cancer: A Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation. Evidence Synthesis No. 124. AHRQ Publication No. 14-05201-EF-1. Rockville, MD: Agency for Healthcare Research and Quality; 2015. Accessed at www.uspreventiveservicestaskforce.org/Page/Document/draft-evidence-review-screening-for-breast-cancer/breast-cancer-screening1 on 22 April 2014.

5. **Melnikow J, Fenton JJ, Whitlock E, Miglioretti D, Weyrich M, Thompson J, et al.** Adjunctive Screening for Breast Cancer in Women With Dense Breasts: A Systematic Review for the U.S. Preventive Services Task Force. Evidence Synthesis No. 126. AHRQ Publication No. 14-05201-EF-2. Rockville, MD: Agency for Healthcare Research and Quality; 2015.

6. **Melnikow J, Fenton JJ, Miglioretti D, Whitlock E, Weyrich M.** Screening for Breast Cancer with Digital Breast Tomosynthesis. Evidence Synthesis No. 125. AHRQ Publication No. 14-05201-EF-2. Rockville, MD: Agency for Healthcare Research and Quality; 2015.

7. **Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, et al; Methods Work Group, Third US Preventive Services Task Force.** Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001;20:21-35. [PMID: 11306229]

8. **U.S. Preventive Services Task Force.** Procedure Manual. Rockville, MD: Agency for Healthcare Research and Quality; 2011.

9. **Nelson HD, Zakher B, Cantor A, Fu R, Griffin J, O'Meara ES, et al.** Risk factors for breast cancer for women aged 40 to 49 years: a systematic review and meta-analysis. Ann Intern Med. 2012;156:635-48. [PMID: 22547473] doi:10.7326/0003-4819-156-9-201205010-00006

10. **Hubbard RA, Kerlikowske K, Flowers CI, Yankaskas BC, Zhu W, Miglioretti DL.** Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. Ann Intern Med. 2011;155:481-92. [PMID: 22007042] doi:10.7326/0003-4819-155-8-201110180-00004

11. **Kerlikowske K, Zhu W, Hubbard RA, Geller B, Dittus K, Braithwaite D, et al; Breast Cancer Surveillance Consortium.** Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. JAMA Intern Med. 2013;173:807-16. [PMID: 23552817] doi:10.1001/jamainternmed.2013.307

12. **Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW.** Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med. 1998;338:1089-96. [PMID: 9545356]

13. **Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M.** The benefits and harms of breast cancer screening: an independent review. Br J Cancer. 2013;108:2205-40. [PMID: 23744281] doi:10.1038/bjc.2013.177

14. **Independent UK Panel on Breast Cancer Screening.** The benefits and harms of breast cancer screening: an independent review. Lancet. 2012;380:1778-86. [PMID: 23117178] doi:10.1016/S0140-6736(12)61611-0

15. **Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge E, Zappa M, et al; EUROSCREEN Working Group.** Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. J Med Screen. 2012;19 Suppl 1:42-56. [PMID: 22972810]

16. **Bleyer A, Welch HG.** Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med. 2012;367:1998-2005. [PMID: 23171096] doi:10.1056/NEJMoa1206809

17. **Coldman A, Phillips N.** Incidence of breast cancer and estimates of overdiagnosis after the initiation of a population-based mammography screening program. CMAJ. 2013;185:E492-8. [PMID: 23754101] doi:10.1503/cmaj.121791

18. **de Gelder R, Fracheboud J, Heijnsdijk EA, den Heeten G, Verbeek AL, Broeders MJ, et al.** Digital mammography screening: weighing reduced mortality against increased overdiagnosis. Prev Med. 2011;53:134-40. [PMID: 21718717] doi:10.1016/j.ypmed.2011.06.009

19. **de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ.** Interpreting overdiagnosis estimates in population-based mammography screening. Epidemiol Rev. 2011;33:111-21. [PMID: 21709144] doi:10.1093/epirev/mxr009

20. **Falk RS, Hofvind S, Skaane P, Haldorsen T.** Overdiagnosis among women attending a population-based mammography screening program. Int J Cancer. 2013;133:705-12. [PMID: 23355313] doi:10.1002/ijc.28052

21. **Hellquist BN, Duffy SW, Nyström L, Jonsson H.** Overdiagnosis in the population-based service screening programme with mammography for women aged 40 to 49 years in Sweden. J Med Screen. 2012;19:14-9. [PMID: 22355181] doi:10.1258/jms.2012.011104

22. **Jørgensen KJ, Zahl PH, Gøtzsche PC.** Overdiagnosis in organised mammography screening in Denmark. A comparative study. BMC Womens Health. 2009;9:36. [PMID: 20028513] doi:10.1186/1472-6874-9-36

23. **Kalager M, Adami HO, Bretthauer M, Tamimi RM.** Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. Ann Intern Med. 2012;156:491-9. [PMID: 22473436] doi:10.7326/0003-4819-156-7-201204030-00005

24. **Martinez-Alonso M, Vilaprinyo E, Marcos-Gragera R, Rue M.** Breast cancer incidence and overdiagnosis in Catalonia (Spain). Breast Cancer Res. 2010;12:R58. [PMID: 20682042] doi:10.1186/bcr2620

25. **Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B.** Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. Cancer Causes Control. 2010;21:275-82. [PMID: 19894130] doi:10.1007/s10552-009-9459-z

26. **Njor SH, Olsen AH, Blichert-Toft M, Schwartz W, Vejborg I, Lynge E.** Overdiagnosis in screening mammography in Denmark: popula-

tion based cohort study. BMJ. 2013;346:f1064. [PMID: 23444414] doi:10.1136/bmj.f1064
27. **Puliti D, Zappa M, Miccinesi G, Falini P, Crocetti E, Paci E.** An estimate of overdiagnosis 15 years after the start of mammographic screening in Florence. Eur J Cancer. 2009;45:3166-71. [PMID: 19879130] doi:10.1016/j.ejca.2009.06.014
28. **Seigneurin A, François O, Labarère J, Oudeville P, Monlong J, Colonna M.** Overdiagnosis from non-progressive cancer detected by screening mammography: stochastic simulation study with calibration to population based registry data. BMJ. 2011;343:d7017. [PMID: 22113564] doi:10.1136/bmj.d7017
29. **Yen AM, Duffy SW, Chen TH, Chen LS, Chiu SY, Fann JC, et al.** Long-term incidence of breast cancer by trial arm in one county of the Swedish Two-County Trial of mammographic screening. Cancer. 2012;118:5728-32. [PMID: 22605639] doi:10.1002/cncr.27580
30. **Zahl PH, Mæhlen J.** Overdiagnosis of breast cancer after 14 years of mammography screening. Tidsskr Nor Laegeforen. 2012;132:414-7. [PMID: 22353833] doi:10.4045/tidsskr.11.0195
31. **Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA.** Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ. 2014;348:g366. [PMID: 24519768] doi:10.1136/bmj.g366
32. **Gunsoy NB, Garcia-Closas M, Moss SM.** Estimating breast cancer mortality reduction and overdiagnosis due to screening for different strategies in the United Kingdom. Br J Cancer. 2014;110:2412-9. [PMID: 24762956] doi:10.1038/bjc.2014.206
33. **Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, et al.** Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. J Med Screen. 2010;17:25-30. [PMID: 20356942] doi:10.1258/jms.2009.009094
34. **Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP.** Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ. 2006;332:689-92. [PMID: 16517548]
35. **Etzioni R, Gulati R, Mallinger L, Mandelblatt J.** Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. Ann Intern Med. 2013;158:831-8. [PMID: 23732716] doi:10.7326/0003-4819-158-11-201306040-00008
36. **Duffy SW, Lynge E, Jonsson H, Ayyaz S, Olsen AH.** Complexities in the estimation of overdiagnosis in breast cancer screening. Br J Cancer. 2008;99:1176-8. [PMID: 18766185] doi:10.1038/sj.bjc.6604638
37. **Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L.** Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. Lancet Oncol. 2007;8:1129-38. [PMID: 18054882]
38. **Miller AB, To T, Baines CJ, Wall C.** Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. J Natl Cancer Inst. 2000;92:1490-9. [PMID: 10995804]
39. **Miller AB, To T, Baines CJ, Wall C.** The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. Ann Intern Med. 2002;137:305-12. [PMID: 12204013] doi:10.7326/0003-4819-137-5_Part_1-200209030-00005
40. **Brett J, Bankhead C, Henderson B, Watson E, Austoker J.** The psychological impact of mammographic screening. A systematic review. Psychooncology. 2005;14:917-38. [PMID: 15786514]
41. **Brewer NT, Salz T, Lillie SE.** Systematic review: the long-term effects of false-positive mammograms. Ann Intern Med. 2007;146:502-10. [PMID: 17404352] doi:10.7326/0003-4819-146-7-200704030-00006
42. **Hafslund B, Nortvedt MW.** Mammography screening from the perspective of quality of life: a review of the literature. Scand J Caring Sci. 2009;23:539-48. [PMID: 19170959] doi:10.1111/j.1471-6712.2008.00634.x
43. **Bond M, Pavey T, Welch K, Cooper C, Garside R, Dean S, et al.** Psychological consequences of false-positive screening mammograms in the UK. Evid Based Med. 2013;18:54-61. [PMID: 22859786] doi:10.1136/eb-2012-100608
44. **Gibson CJ, Weiss J, Goodrich M, Onega T.** False-positive mammography and depressed mood in a screening population: findings from the New Hampshire Mammography Network. J Public Health (Oxf). 2009;31:554-60. [PMID: 19574274] doi:10.1093/pubmed/fdp064
45. **Keyzer-Dekker CM, De Vries J, van Esch L, Ernst MF, Nieuwenhuijzen GA, Roukema JA, et al.** Anxiety after an abnormal screening mammogram is a serious problem. Breast. 2012;21:83-8. [PMID: 21924905] doi:10.1016/j.breast.2011.08.137
46. **Brodersen J, Siersma VD.** Long-term psychosocial consequences of false-positive screening mammography. Ann Fam Med. 2013;11:106-15. [PMID: 23508596] doi:10.1370/afm.1466
47. **Tosteson AN, Fryback DG, Hammond CS, Hanna LG, Grove MR, Brown M, et al.** Consequences of false-positive screening mammograms. JAMA Intern Med. 2014;174:954-61. [PMID: 24756610] doi:10.1001/jamainternmed.2014.981
48. **Espasa R, Murta-Nascimento C, Bayés R, Sala M, Casamitjana M, Macià F, et al.** The psychological impact of a false-positive screening mammogram in Barcelona. J Cancer Educ. 2012;27:780-5. [PMID: 22477233] doi:10.1007/s13187-012-0349-9
49. **Schou Bredal I, Kåresen R, Skaane P, Engelstad KS, Ekeberg Ø.** Recall mammography and psychological distress. Eur J Cancer. 2013;49:805-11. [PMID: 23021930] doi:10.1016/j.ejca.2012.09.001
50. **Hafslund B, Espehaug B, Nortvedt MW.** Effects of false-positive results in a breast screening program on anxiety, depression and health-related quality of life. Cancer Nurs. 2012;35:E26-34. [PMID: 22067696] doi:10.1097/NCC.0b013e3182341ddb
51. **Fitzpatrick P, Fleming P, O'Neill S, Kiernan D, Mooney T.** False-positive mammographic screening: factors influencing re-attendance over a decade of screening. J Med Screen. 2011;18:30-3. [PMID: 21536814] doi:10.1258/jms.2010.010104
52. **Klompenhouwer EG, Duijm LE, Voogd AC, den Heeten GJ, Strobbe LJ, Louwman MW, et al.** Re-attendance at biennial screening mammography following a repeated false positive recall. Breast Cancer Res Treat. 2014;145:429-37. [PMID: 24748569] doi:10.1007/s10549-014-2959-x
53. **Maxwell AJ, Beattie C, Lavelle J, Lyburn I, Sinnatamby R, Garnett S, et al.** The effect of false positive breast screening examinations on subsequent attendance: retrospective cohort study. J Med Screen. 2013;20:91-8. [PMID: 24009091] doi:10.1177/0969141313499147
54. **Lowe JB, Balanda KP, Del Mar C, Hawes E.** Psychologic distress in women with abnormal findings in mass mammography screening. Cancer. 1999;85:1114-8. [PMID: 10091796]
55. **Hislop TG, Harris SR, Jackson J, Thorne SE, Rousseau EJ, Coldman AJ, et al.** Satisfaction and anxiety for women during investigation of an abnormal screening mammogram. Breast Cancer Res Treat. 2002;76:245-54. [PMID: 12462385]
56. **Brett J, Austoker J.** Women who are recalled for further investigation for breast screening: psychological consequences 3 years after recall and factors affecting re-attendance. J Public Health Med. 2001;23:292-300. [PMID: 11873891]
57. **Ong G, Austoker J, Brett J.** Breast screening: adverse psychological consequences one month after placing women on early recall because of a diagnostic uncertainty. A multicentre study. J Med Screen. 1997;4:158-68. [PMID: 9368874]
58. **Aro AR, Pilvikki Absetz S, van Elderen TM, van der Ploeg E, van der Kamp LJ.** False-positive findings in mammography screening induces short-term distress-breast cancer-specific concern prevails longer. Eur J Cancer. 2000;36:1089-97. [PMID: 10854941]
59. **Brett J, Austoker J, Ong G.** Do women who undergo further investigation for breast screening suffer adverse psychological consequences? A multi-centre follow-up study comparing different breast screening result groups five months after their last breast screening appointment. J Public Health Med. 1998;20:396-403. [PMID: 9923945]
60. **Lerman C, Trock B, Rimer BK, Boyce A, Jepson C, Engstrom PF.** Psychological and behavioral implications of abnormal mammo-

grams. Ann Intern Med. 1991;114:657-61. [PMID: 2003712] doi:10.7326/0003-4819-114-8-657
61. **Lampic C, Thurfjell E, Sjödén PO.** The influence of a false-positive mammogram on a woman's subsequent behaviour for detecting breast cancer. Eur J Cancer. 2003;39:1730-7. [PMID: 12888368]
62. **Cockburn J, Staples M, Hurley SF, De Luise T.** Psychological consequences of screening mammography. J Med Screen. 1994;1:7-12. [PMID: 8790480]
63. **Ekeberg Ø, Skjauff H, Kåresen R.** Screening for breast cancer is associated with a low degree of psychological distress. Breast. 2001;10:20-4. [PMID: 14965553]
64. **Gilbert FJ, Cordiner CM, Affleck IR, Hood DB, Mathieson D, Walker LG.** Breast screening: the psychological sequelae of false-positive recall in women with and without a family history of breast cancer. Eur J Cancer. 1998;34:2010-4. [PMID: 10070302]
65. **Gram IT, Lund E, Slenker SE.** Quality of life following a false positive mammogram. Br J Cancer. 1990;62:1018-22. [PMID: 2257206]
66. **Lampic C, Thurfjell E, Bergh J, Sjödén PO.** Short- and long-term anxiety and depression in women recalled after breast cancer screening. Eur J Cancer. 2001;37:463-9. [PMID: 11267855]
67. **Lidbrink E, Levi L, Pettersson I, Rosendahl I, Rutqvist LE, de la Torre B, et al.** Single-view screening mammography: psychological, endocrine and immunological effects of recalling for a complete three-view examination. Eur J Cancer. 1995;31A:932-3. [PMID: 7646925]
68. **Sandin B, Chorot P, Valiente RM, Lostao L, Santed MA.** Adverse psychological effects in women attending a second-stage breast cancer screening. J Psychosom Res. 2002;52:303-9. [PMID: 12023127]
69. **McCann J, Stockton D, Godward S.** Impact of false-positive mammography on subsequent screening attendance and risk of cancer. Breast Cancer Res. 2002;4:R11. [PMID: 12223128]
70. **O'Sullivan I, Sutton S, Dixon S, Perry N.** False positive results do not have a negative effect on reattendance for subsequent breast screening. J Med Screen. 2001;8:145-8. [PMID: 11678554]
71. **Orton M, Fitzpatrick R, Fuller A, Mant D, Mlynek C, Thorogood M.** Factors affecting women's response to an invitation to attend for a second breast cancer screening examination. Br J Gen Pract. 1991;41:320-2. [PMID: 1777275]
72. **Meldrum P, Turnbull D, Dobson HM, Colquhoun C, Gilmour WH, McIlwaine GM.** Tailored written invitations for second round breast cancer screening: a randomised controlled trial. J Med Screen. 1994;1:245-8. [PMID: 8790529]
73. **Armstrong K, Moye E, Williams S, Berlin JA, Reynolds EE.** Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. Ann Intern Med. 2007;146:516-26. [PMID: 17404354] doi:10.7326/0003-4819-146-7-200704030-00008
74. **Whelehan P, Evans A, Wells M, Macgillivray S.** The effect of mammography pain on repeat participation in breast cancer screening: a systematic review. Breast. 2013;22:389-94. [PMID: 23541681] doi:10.1016/j.breast.2013.03.003
75. **Miller D, Livingstone V, Herbison P.** Interventions for relieving the pain and discomfort of screening mammography. Cochrane Database Syst Rev. 2008:CD002942. [PMID: 18254010] doi:10.1002/14651858.CD002942.pub2
76. **Drossaert CH, Boer H, Seydel ER.** Health education to improve repeat participation in the Dutch breast cancer screening programme: evaluation of a leaflet tailored to previous participants. Patient Educ Couns. 1996;28:121-31. [PMID: 8852086]
77. **Drossaert CH, Boer H, Seydel ER.** Monitoring women's experiences during three rounds of breast cancer screening: results from a longitudinal study. J Med Screen. 2002;9:168-75. [PMID: 12518007]
78. **Rutter DR, Calnan M, Field S, Vaile MSB.** Predicting reattendance in the second round of the UK National Breast Screening Programme: a prospective 3-year longitudinal analysis. Breast. 1997;6:120-5. doi:10.1016/S0960-9776(97)90551-8
79. **Aro AR, Absetz-Ylöstalo P, Eerola T, Pamilo M, Lönnqvist J.** Pain and discomfort during mammography. Eur J Cancer. 1996;32A:1674-9. [PMID: 8983273]
80. **Bennett IC, Robert DA, Osborne JM, Baker CA.** Discomfort during mammography: a survey of women attending a breast screening center. Breast Dis. 1994;7:35-41.
81. **Drossaert CH, Boer H, Seydel ER.** Does mammographic screening and a negative result affect attitudes towards future breast screening? J Med Screen. 2001;8:204-12. [PMID: 11743037]
82. **Dullum JR, Lewis EC, Mayer JA.** Rates and correlates of discomfort associated with mammography. Radiology. 2000;214:547-52. [PMID: 10671609]
83. **Papas MA, Klassen AC.** Pain and discomfort associated with mammography among urban low-income African-American women. J Community Health. 2005;30:253-67. [PMID: 15989208]
84. **Alimoglu E, Alimoglu MK, Kabaalioglu A, Ceken K, Apaydin A, Lüleci E.** [Mammography-related pain and anxiety]. Tani Girisim Radyol. 2004;10:213-7. [PMID: 15470624]
85. **Shrestha S, Poulos A.** The effect of verbal information on the experience of discomfort in mammography. Radiography. 2001;7:271-7. doi:10.1053/radi.2001.0344
86. **Sjölin M, Maerker M.** [Effects of information on and reflection of women's pain experience during mammography]. Vard Nord Utveckl Forsk. 1994;14:11-5. [PMID: 7941212]
87. **Kornguth PJ, Rimer BK, Conaway MR, Sullivan DC, Catoe KE, Stout AL, et al.** Impact of patient-controlled compression on the mammography experience. Radiology. 1993;186:99-102. [PMID: 8416595]
88. **Poulos A, Rickard M.** Compression in mammography and the perception of discomfort. Australas Radiol. 1997;41:247-52. [PMID: 9293675]
89. **Lambertz CK.** Pre-medication with Acetaminophen for Screening Mammography. Spokane, WA: Gonzaga Univ; 1998.
90. **Dibble SL, Israel J, Nussey B, Sayre JW, Brenner RJ, Sickles EA.** Mammography with breast cushions. Womens Health Issues. 2005;15:55-63. [PMID: 15767195]
91. **Hendrick RE.** Radiation doses and cancer risks from breast imaging studies. Radiology. 2010;257:246-53. [PMID: 20736332] doi:10.1148/radiol.10100570
92. **Yaffe MJ, Mainprize JG.** Risk of radiation-induced breast cancer from mammographic screening. Radiology. 2011;258:98-105. [PMID: 21081671] doi:10.1148/radiol.10100655
93. **Friedewald SM, Rafferty EA, Rose SL, Durand MA, Plecha DM, Greenberg JS, et al.** Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA. 2014;311:2499-507. [PMID: 25058084] doi:10.1001/jama.2014.6095
94. **Rose SL, Tidwell AL, Bujnoch LJ, Kushwaha AC, Nordmann AS, Sexton R Jr.** Implementation of breast tomosynthesis in a routine screening practice: an observational study. AJR Am J Roentgenol. 2013;200:1401-8. [PMID: 23701081] doi:10.2214/AJR.12.9672
95. **Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, et al.** Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. Lancet Oncol. 2013;14:583-9. [PMID: 23623721] doi:10.1016/S1470-2045(13)70134-7
96. **Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U, et al.** Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. Radiology. 2013;267:47-56. [PMID: 23297332] doi:10.1148/radiol.12121373
97. **Haas BM, Kalra V, Geisel J, Raghu M, Durand M, Philpotts LE.** Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening. Radiology. 2013;269:694-700. [PMID: 23901124] doi:10.1148/radiol.13130307
98. **Chiarelli AM, Majpruz V, Brown P, Thériault M, Shumak R, Mai V.** The contribution of clinical breast examination to the accuracy of breast screening. J Natl Cancer Inst. 2009;101:1236-43. [PMID: 19720967] doi:10.1093/jnci/djp241

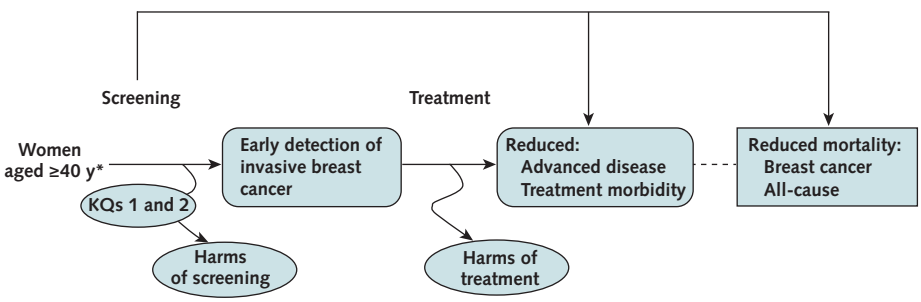
Current Author Addresses: Drs. Nelson, Cantor, and Humphrey; Ms. Pappas; Ms. Griffin; and Ms. Daeges: Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Mail Code BICC, Portland, OR 97239.

Author Contributions: Conception and design: H.D. Nelson, M. Pappas, A. Cantor, L. Humphrey.
Analysis and interpretation of the data: H.D. Nelson, M. Pappas, A. Cantor, J. Griffin, M. Daeges, L. Humphrey.
Drafting of the article: H.D. Nelson, M. Pappas, A. Cantor, L. Humphrey.
Critical revision of the article for important intellectual content: H.D. Nelson, A. Cantor, L. Humphrey.
Final approval of the article: H.D. Nelson, M. Pappas, A. Cantor, J. Griffin, M. Daeges, L. Humphrey.
Provision of study materials or patients: H.D. Nelson, M. Daeges.
Obtaining of funding: H.D. Nelson.
Administrative, technical, or logistic support: H.D. Nelson, M. Pappas, A. Cantor, M. Daeges.
Collection and assembly of data: H.D. Nelson, M. Pappas, A. Cantor, J. Griffin, M. Daeges, L. Humphrey.

Web-Only References

99. **de Koning HJ, Draisma G, Fracheboud J, de Bruijn A.** Overdiagnosis and overtreatment of breast cancer: microsimulation modelling estimates based on observed screen and clinical data. Breast Cancer Res. 2006;8:202. [PMID: 16524452]
100. **Duffy SW, Agbaje O, Tabar L, Vitak B, Bjurstam N, Björneld L, et al.** Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. Breast Cancer Res. 2005;7:258-65. [PMID: 16457701]
101. **Olsen AH, Agbaje OF, Myles JP, Lynge E, Duffy SW.** Overdiagnosis, sojourn time, and sensitivity in the Copenhagen mammography screening program. Breast J. 2006;12:338-42. [PMID: 16848843]
102. **Paci E, Warwick J, Falini P, Duffy SW.** Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? J Med Screen. 2004;11:23-7. [PMID: 15006110]
103. **Paci E, Miccinesi G, Puliti D, Baldazzi P, De Lisi V, Falcini F, et al.** Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy. Breast Cancer Res. 2006;8:R68. [PMID: 17147789]
104. **Yen MF, Tabár L, Vitak B, Smith RA, Chen HH, Duffy SW.** Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. Eur J Cancer. 2003;39:1746-54. [PMID: 12888370]
105. **Zahl PH.** Overdiagnosis of breast cancer in Denmark [Letter]. Br J Cancer. 2004;90:1686. [PMID: 15083204]

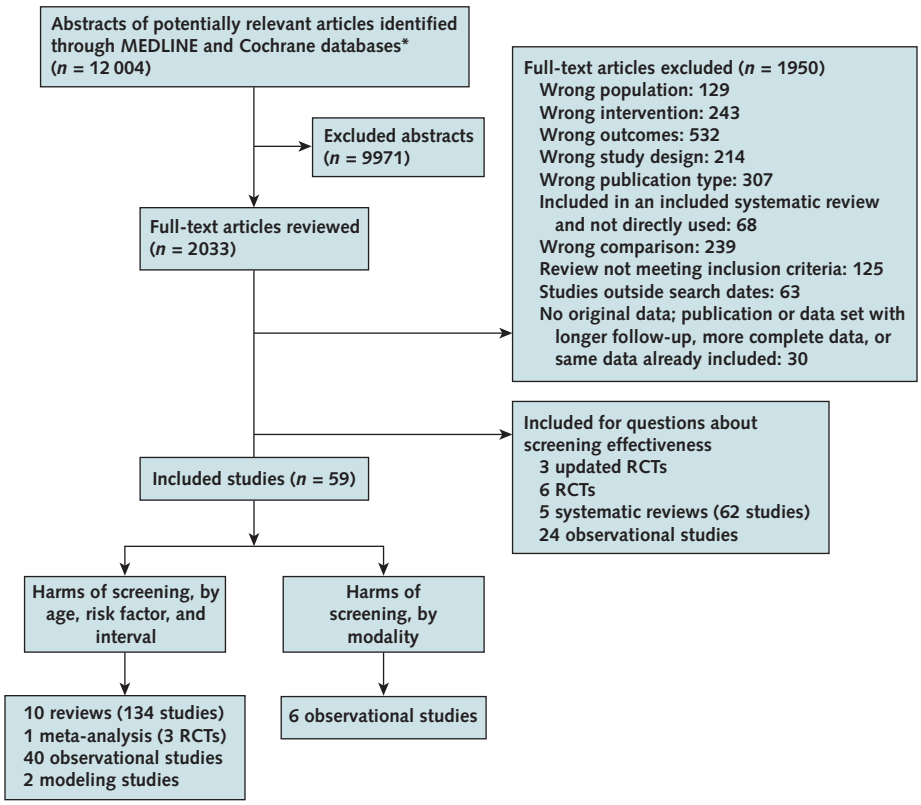
Appendix Figure 1. Analytic framework and key questions.



Key Questions:
For women aged ≥40 y*:
1. What are the harms† of routine mammography screening, and how do they differ by age, risk factor‡, and screening interval?
2. How do the harms† of routine breast cancer screening vary by screening modality§?

KQ = key question.
* Excludes women with preexisting breast cancer; clinically significant *BRCA1* or *BRCA2* mutations, Li-Fraumeni syndrome, Cowden syndrome, hereditary diffuse gastric cancer, or other familial breast cancer syndrome; high-risk lesions (ductal carcinoma in situ, lobular carcinoma in situ, atypical ductal hyperplasia, or atypical lobular hyperplasia); or previous large doses of chest radiation (≥20 Gy) before age 30 y.
† False-positive and false-negative mammography results, biopsy recommendations due to false-positive mammography results, overdiagnosis and resulting overtreatment, anxiety, pain, and radiation exposure.
‡ 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.
§ Mammography (film, digital, or tomosynthesis), magnetic resonance imaging, ultrasonography, and clinical breast examination (alone or in combination).

Appendix Figure 2. Summary of evidence search and selection.



RCT = randomized, controlled trial.
* Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews.

Appendix Table 1. U.S. Studies of Cumulative False-Positive Mammography and Biopsy Results

Author, Year (Reference)	Study Design	Population	Age, *y*	Participants, *n*	Study Years	Comparison	Outcome Measures	Results
New studies								
Hubbard et al, 2011 (10)	Postintervention series	U.S., 7 mammography registries in the BCSC	40-59	169 456	1994-2006	Annual vs. biennial screening by age	FP results (no diagnosis of invasive carcinoma or DCIS within 1 y of screening or before the next screening mammogram); recalls (BI-RADS 0, 3, 4, 5)	Cumulative probability of FP mammography after 10 y, % (95% CI) Age 40: annual, 61.3 (59.4 to 63.1); biennial, 41.6 (40.6 to 42.5) Age 50: annual, 61.3 (58.0 to 64.7); biennial, 42.0 (40.4 to 43.7) Cumulative probability of FP biopsy after 10 y, % (95% CI) Age 40: annual, 7.0 (6.1 to 7.8); biennial, 4.8 (4.4 to 5.2) Age 50: annual, 9.4 (7.4 to 11.5); biennial, 6.4 (5.6 to 7.2)
Kerlikowske et al, 2013 (11)	Postintervention series	U.S., 7 mammography registries in the BCSC	40-74	11 474 with breast cancer, 922 624 without	1994-2008	Annual vs. biennial vs. triennial screening by age, breast density, and menopausal hormone therapy	FP results (no diagnosis of invasive carcinoma or DCIS within 1 y of screening or before the next screening mammogram); recalls (BI-RADS 0, 3, 4, 5)	Cumulative probability of FP mammography after 10 y, by breast density*, % (95% CI) Age 40-49: annual: 36 (34 to 38); 60 (59 to 61); 69 (68 to 70); 66 (64 to 67); biennial: 21 (20 to 22); 39 (38 to 39); 46 (46 to 47); 43 (42 to 44); triennial: 14 (13 to 15); 27 (26 to 27); 33 (31 to 34); 33 (32 to 34) Age 50-74: annual: 30 (29 to 31); 50 (49 to 51); 60 (59 to 61); 59 (57 to 60); biennial: 17 (17 to 18); 31 (30 to 31); 39 (38 to 39); 38 (37 to 38); triennial: 12 (12 to 13); 22 (21 to 22); 28 (28 to 29); 27 (26 to 28) Cumulative probability of FP biopsy after 10 y, by breast density*, % (95% CI) Age 40-49: annual: 6 (5 to 7); 9 (8 to 10); 12 (11 to 14); 12 (11 to 14); biennial: 3 (2 to 3); 5 (4 to 5); 7 (6 to 7); 7 (6 to 7); triennial: 2 (2 to 2); 3 (3 to 4); 4 (3 to 4); 3 (2 to 4) Age 50-74: annual: 5 (5 to 6); 8 (8 to 9); 11 (10 to 12); 11 (10 to 12); biennial: 3 (3 to 3); 5 (4 to 5); 6 (6 to 7); 6 (6 to 7); triennial: 2 (2 to 2); 3 (3 to 4); 5 (4 to 5); 5 (4 to 5) Highest cumulative rates of FP mammography (66% to 69%) or biopsy (12% to 14%): annual mammography; extremely or heterogeneously dense breasts; age 40-49; used combined hormone therapy
2009 review								
Elmore et al, 1998 (12)	Postintervention series	U.S., randomly sampled patients from 11 health centers in an HMO	40-69	NR	1983-1995	Annual vs. biennial screening	FP results (not a true positive = breast cancer diagnosed on the basis of pathologic findings within 1 y of mammography)	Cumulative risk for at least one FP after 10 screening mammograms, % (95% CI) Age 40-49: 56 (39.5 to 75.8) Age 50-59: 47 (37.8 to 63.0) Overall: 49 (40.3 to 64.1) Cumulative risk for FP biopsy, % (95% CI) Overall: 19 (9.8 to 41.2)

BCSC = Breast Cancer Surveillance Consortium; BI-RADS = Breast Imaging Reporting and Data System; DCIS = ductal carcinoma in situ; FP = false-positive; NR = not reported.
* Almost entirely fat, scattered fibroglandular densities, heterogenously dense, or extremely dense.

Appendix Table 2. Studies of Overdiagnosis With Breast Cancer Screening

Author, Year (Reference)	Age, *y*	Study Years	Data Source	Comparison Groups	Approach, Lead-Time Adjustment	Overdiagnosis Measures	Rates of Invasive Cancer + DCIS	Rates of Invasive Cancer	Rates of DCIS
New studies									
Bleyer and Welch, 2012 (16)	≥40	1976-2008	SEER; United States	Population before vs. after widespread screening	EI; no adjustment	Change in incidence before and after introduction of screening with 3 estimates of baseline incidence Best guess: incidence increases 0.25% annually Extreme: incidence increases 0.50% annually Very extreme: using highest observed incidence, assume a 0.50% incidence increase	Best guess: 31% Extreme: 26% Very extreme: 22%	NR	NR
Coldman and Phillips, 2013 (17)	40-89	1970-2009	Breast cancer registry; Canada	Population before vs. after widespread screening	EI; compensatory drop	Participation estimate: cumulative incidence with active screening vs. never screened or nonactive screening Population estimate: observed vs. expected population cumulative incidence in 2005-2009	Participation estimate: 17.3% Population estimate: 6.7%	Participation estimate: 5.4% Population estimate: −0.7%	NR
de Gelder et al, 2011 (18)	49-74	2004-2006	Screening program (biennial); the Netherlands	Modeled incidence of screening vs. predicted incidence without screening	LT; statistical adjustment; preclinical DCIS: mean 5.2 y; preclinical invasive: 2.6 y	Microsimulation analysis (digital mammography) Baseline model: 18% are screen-detectable preclinical DCIS; 11% progress to invasive cancer, 5% are clinically diagnosed, 2% regress Progressive model: all tumors have preclinical screen-detectable DCIS stage and none regress; 96% invasive with no screening, 4% are clinically diagnosed Nonprogressive model: no preclinical screen-detected DCIS, majority regress, 2% are clinically diagnosed	Baseline model: 2.5% all cases; 8.2% screen-detected Progressive model: 1.4% all cases; 5.0% screen-detected Nonprogressive model: 7.7% all cases; 25.2% screen-detected	NR	NR
de Gelder et al, 2011* (19)	0-69; 0-74	1990-1998; 1998-2007	Screening program (biennial); the Netherlands	Modeled incidence of screening vs. predicted incidence without screening	LT; compensatory drop; mean 2.6 y	Microsimulation screening analysis; excess cancers minus deficit cancers divided by the total number of breast cancers in the absence of screening in women 0-100 y	1-y estimates 1990-1998: 1.0%; 6.1%; 9.1%; 11.4%; 10.0%; 9.4%; 8.8%; 5.6% 1-y estimates 1998-2007: 4.9%; 10.0%; 7.4%; 4.7%; 4.7%; 4.9%; 4.3%; 4.4%; 2.8%	NR	NR
Duffy et al, 2010 (33)	50-69	1977-1998; 1974-2003	Swedish Two-County Trial; U.K. National Breast Screening Program	Active vs. passive screening; population before vs. after widespread screening	EI; compensatory drop	Swedish Trial: Estimated expected incidence trends in the prescreening period vs. observed cases, adjusted for prevalence peak U.K. Program: Observed cases of breast cancer, minus any deficit in ages 65-69 or ≥70 y	Overall: 4% to 7% Swedish Trial: 4.3 cases per 1000 women screened for 20 y U.K. Program: 2.3 cases per 1000 women screened for 20 y	NR	NR
Falk et al, 2013† (20)	50-69	1995-2009	Norwegian Breast Cancer Screening Program (biennial)	Women screened vs. those never invited or did not attend screening	EI; compensatory drop	Women attending screening adjusted for adherence to screening vs. 3 reference rates: 40-year-olds 1993-1995 Observed rates of invasive breast cancer 1980-1984 Cohort of women born 1903-1907	16.5%; 16.3%; 13.9%	11.3%; 11.2%; 9.6%	NR
Gunsoy et al, 2014 (32)	40-73	1971-2010	Data from various sources in the U.K.	Women screened vs. not screened	Multiple statistical adjustments	Markov model of the difference between cumulative incidence of invasive + DCIS with denominators: Cases diagnosed in absence of screening age 40-85 Cases diagnosed in screening period Screen-detected breast cancers	All cases: 4.3 to 8.9% Screening period: 6.7% to 10.1% Screen-detected: 11.8% to 13.5% Highest rates with frequent screening	NR	NR

Continued on following page

Appendix Table 2—Continued

Author, Year (Reference)	Age, y	Study Years	Data Source	Comparison Groups	Approach, Lead-Time Adjustment	Overdiagnosis Measures	Rates of Invasive Cancer + DCIS	Rates of Invasive Cancer	Rates of DCIS
Hellquist et al, 2012 (21)	40-49	1986-2005	Screening for Young Women Trial; Sweden	Population in areas with vs. without screening	EI; statistical adjustment; up to 1.5 y	Incidence in screening group vs. controls Corrected for prescreening difference, prevalence peak bias (excluded prevalence screen data), trend bias (change in incidence per year of age)	Rate ratio: 1.01 (95% CI 0.94 to 1.08)	Rate ratio: 0.95 (95% CI 0.88 to 1.01)	NR
Jørgensen et al, 2009‡ (22)	50-69	1991-2003 vs. 1971-1990	Screening program; Copenhagen and Funen, Denmark	Population in areas with (1991-2003) vs. without (1971-1990) screening	EI; compensatory drop	Ratio of incidence between screened and nonscreened areas for the screened age group	33%	NR	NR
Kalager et al, 2012§ (23)	50-69	1996-2005	Norwegian Breast Cancer Screening Program (biennial)	Population in areas with vs. without screening	EI; compensatory drop; approach 1: 10-y lead time; approach 2: 5 or 2 y	Approach 1: Incidence rates in the screening and nonscreening groups for women aged 50-79 y Approach 2: Excluded all cases of cancer detected in the first screening round, compares incidence in screened women vs. women 2-5 y older	NR	Approach 1: entire country: 25%, region 1: 18% Approach 2: 5-y lead time: 15%, 2-y lead time: 20%	NR
Martinez-Alonso et al, 2010 (24)	40-69	1980-2004	Cancer registry; Catalonia, Spain	Modeled pre vs. post screening incidence	EI; statistical adjustments	Probabilistic model for birth cohorts: 1935, 1940, 1945, 1950; observed vs. expected cumulative incidence	NR	1935: 0.4% 1940: 23.3% 1945: 30.6% 1950: 46.6%	NR
Miller et al, 2014 (31)	40-59	1980-1985	Canadian National Breast Screening Study	Randomized trial; screening vs. usual care	EI; none	Excess of breast cancer cases in mammography group vs. control group of trial	NR	22% of screen-detected cancer	NR
Morrell et al, 2010 (25)	50-69	1999-2001	Screening program (biennial); Australia	Screened vs. unscreened age group or before screening implementation	EI; statistical adjustment; 2- or 5-y lead times	Observed annual incidence minus expected annual incidence divided by expected annual incidence Interpolation approach: incidence in unscreened women (≤40 or ≥80) modeled by 5-y age group Extrapolation approach: incidence for the period before the introduction of screening modeled for all 5-y age groups and extrapolated to 1999-2001	NR	Interpolation: 2-y: 51%; 5-y: 42%; Extrapolation: 2-y: 36%, 5-y: 30% Rates higher for 50-59 vs. 60-69	NR
Njor et al, 2013 (26)	56-70	1991-2005	Screening program; Copenhagen and Funen, Denmark	Population in areas with vs. without screening	EI; compensatory drop	Cumulative incidence in screened population vs. expected incidence in unscreened counties	≥8 y follow-up: Copenhagen, 3% (−14% to 25%), Funen, 0.7% (−9% to 12%)	NR	NR
Puliti et al, 2009 (27)	60-69	1990-NR	Screening program; Florence, Italy	Screening vs. prescreening	EI; compensatory drop	Ratio of cumulative incidence of breast cancer in the invited group to those in the noninvited group at least 5 y after last screening, assuming 1.2% annual trend in prescreening incidence	Rate ratio: 1.01 (95% CI 0.95 to 1.07)	Rate ratio: 0.99 (95% CI 0.94 to 1.05)	NR
Seigneurin et al, 2011 (28)	50-69	1991-2006	Cancer registry; Isere, France	Modeled screening incidence	LT; statistical adjustment, 2-4 y	Stochastic simulation model, driven by all-cause mortality, lifetime probability of breast cancer, natural course of breast cancer, and cancer detection; adjusted for sojourn time	NR	All diagnosed cancers: 1.5%, screen-detected: 3.3%	All diagnosed cancers: 28.0%, screen-detected: 31.9%
Yen et al, 2012 (29)	40-74	1977-2005	Swedish Two-County Trial; data from one county only (Dalarna)	Active screening vs. passive screening	EI; compensatory drop	Cumulative incidence in active screening vs. usual care groups	Relative risk: 1.00 (95% CI 0.92 to 1.08)	Relative risk: 0.99 (95% CI 0.88 to 1.55)	Relative risk: 1.17 (95% CI 0.88 to 1.55)
Zahl and Mæhlen, 2012 (30)	40-79	1991-2009	Norway Cancer Registry	Screening vs. postscreening	EI; compensatory drop	Define overdiagnosis as increase in number of cancer diagnoses among those who are invited for screening and the reduction in the number of diagnoses among those no longer invited	~50%	NR	NR

Continued on following page

Appendix Table 2—Continued

Author, Year (Reference)	Age, y	Study Years	Data Source	Comparison Groups	Approach, Lead-Time Adjustment	Overdiagnosis Measures	Rates of Invasive Cancer + DCIS	Rates of Invasive Cancer	Rates of DCIS
2009 review									
de Koning et al, 2006 (99)	50-74	1989-2001	National data from the Netherlands	Screening vs. nonscreening (biennial)	Statistical adjustments; assumptions of DCIS progression	Microsimulation model	3% in screened population; 8% screen-detected	NR	NR
Duffy et al, 2005 (100)	40-74	1977-1985	Swedish Two-County Trial	Active vs. passive screening	Lead-time statistical adjustments	Markov multistate model	1% in screened population	NR	NR
	39-59	1982-1996	Gothenburg trial	Screening vs. no screening	Lead-time statistical adjustments	Markov multistate model	2% in screened population	NR	NR
Olsen et al, 2006 (101)	50-71	1991-1996	Copenhagen, Denmark; screening program (biennial)	Incidence in screened women	Statistical adjustments	Chronic disease statistical model of screen-detected overdiagnosis	Prevalence: 7.8% Incidence: 0.5%	NR	NR
Paci et al, 2004 (102)	50-69	1985-1999	Florence, Italy; screening program	Incidence in screening vs. prescreening	EI; corrected for lead time	Observed/expected cases	5%	2%	3%
Paci et al, 2006 (103)	50-74	1986-2001	Italy; screening program	Prescreening incidence	EI; corrected for lead time	Observed/expected cases	4.6%; range −0.6% to 9.7% varies by age (highest in 50-54 and 65-74)	3.2%	1.4%
Yen et al, 2003 (104)	40-69	NR	Swedish Two-County Trial, United Kingdom, the Netherlands, Australia, New York	Screening vs. no screening	LT; statistical adjustment	6-state Markov model	NR	NR	Prevalence: 37% Incidence: 4%
	40-69	NR	Swedish Two-County Trial	Screening vs. no screening	LT; statistical adjustment	6-state Markov model	NR	NR	40-49: 19%, 3% 50-59: 23%, 4% 60-69: 46%, 6%
Zackrisson et al, 2006 (34)	55-69	1978-1986	Malmö trial	Randomized screening vs. no screening	EI; compensatory drop	Comparison of incidence in screened vs. unscreened	10% of incidence in control group	7%	3%
Zahl, 2004 (105)	50-69	1971-2000	Norway and Sweden	Prescreening incidence	EI; compensatory drop	Changes in age-specific incidence rates associated with the introduction of screening programs	NR	30% of incidence in screened population	NR

DCIS = ductal carcinoma in situ; EI = excess incidence approach; LT = lead-time approach; NR = not reported; SEER = Surveillance, Epidemiology, and End Results Program.
* Additional 6 model estimates for each year are published in this paper to show that the range of estimates varies by selection of the denominator.
† Population overlap with Kalager and colleagues (23).
‡ Same Copenhagen population as Olsen and colleagues (101).
§ Population overlap with Falk and colleagues (20).

Appendix Table 3. Systematic Reviews of Pain With Mammography

Author, Year (Reference)	Inclusion Criteria	Searches	Studies, *n* (Designs); Participants, *n*	Methods	Results	Quality Rating	Limitations
New studies							
Whelehan et al, 2013 (74)	Studies of pain or discomfort of screening mammography and reattendance	MEDLINE, EMBASE, PsycINFO, CINAHL, ASSIA, Cochrane Database of Systematic Reviews, Sociological Abstracts, SSCI, SCI, and NHS online literature database; October 2012	20 (most cross-sectional surveys); causation (*n* = 5741); association (*n* = NR)	Quality based on individual factors*; studies combined separately for causation vs. association	Causation (7 studies); response rates: 32-79% Actual nonreattendance indicating pain as the reason (5 studies): 11-46% Intended future nonreattendance due to pain (2 studies): 2.7% and 17.5% Association (15 studies) Actual reattendance (10 studies): no difference between women who experienced pain vs. no pain (RR 1.38; 95% CI 0.94 to 2.02; 5 studies); higher pain scores in non reattenders vs. reattenders in 2 of 3 studies ($P = 0.001$ and $P < 0.05$) Intended reattendance (5 studies): no differences (3 studies), less intent for women with pain (2 studies) with OR 0.61 (95% CI 0.38 to 0.98) in one study	Fair	Unclear how study quality was used to formulate conclusions; did not describe characteristics of all included studies; did not assess publication bias
2009 review							
Armstrong et al, 2007 (73)	Studies of risks of screening mammography for women in their 40s	MEDLINE, PreMEDLINE, and the Cochrane Central Register of Controlled Trials; May 2005	22 (3 RCTs, 5 prospective cohort, 1 retrospective cohort, 13 cross-sectional); 13 008	Centre for Evidence-based Medicine criteria; based on study design and rates of attrition; methods of synthesis not described	Prevalence of pain from mammography varied from 28-77% Degree of pain was associated with stage of menstrual cycle (3 studies), anxiety (2 studies), and premammography anticipation of pain (4 studies)	Fair	No synthesis of data; unclear how study quality was used to formulate conclusions; study designs not prespecified; did not assess publication bias
Miller et al, 2008 (75)	RCTs of interventions that reduce or relieve the pain and discomfort of screening mammography	MEDLINE, EMBASE, CINAHL, and Cochrane Breast Cancer Specialised Register; 2006	7 (RCT); 1771	Based on generation and concealment of allocation sequence, comparability of groups at baseline, intention-to-treat analysis, and double-blinding after allocation	Information provided before mammography vs. usual care (3 trials): 44% vs. 24% ($P = 0.009$) experienced less discomfort than expected with verbal information (1 trial) Pain scores were lower with written information in 1 trial (mean VAS score 16.5 vs. 24.5; $P < 0.05$), but no differences were found in another trial Breast compression strategies (2 trials): Participant vs. technologist compression indicated 57% felt no difference in discomfort, 31% less, 13% more No difference with normal vs. 1 second of reduced compression Premedication (1 study): acetaminophen vs. none (mean VAS scores 23.7 vs. 22.8; $P = 0.896$) Breast cushion (1 study): reduced pain for cushion vs. no cushion (mean VAS pain score 20.34 vs. 34.94; $P < 0.0001$)	Good	Did not assess publication bias

ASSIA = Applied Social Sciences Index and Abstracts; NHS = National Health Service; NR = not reported; OR = odds ratio; RCT = randomized, controlled trial; RR = risk ratio; SCI = Science Citation Index; SSCI = Social Sciences Citation Index; VAS = visual analogue scale.

* Includes whether intended or actual reattendance was measured, survey response rate/participation rate, measures of pain or discomfort, consistency of the timing of outcome measurement, quality of statistical analysis, and robustness of ascertaining reattendance rate.

Appendix Table 4. Models of Radiation Exposure With Screening, Breast Cancer Incidence, and Death

Author, Year (Reference)	Study Design	Population	Age, *y*	Method	Outcome Measures	Results
Hendrick, 2010 (91)	Modeling study	U.S.-based sources	40 to 80	Theoretical estimates based on long-term follow-up of acute exposures to higher levels of ionizing radiation and a linear no-threshold extrapolation of risks at low doses. Model assumes 3.7 mGy to 4.7 mGy per examination.	Breast cancer cases and mortality	LAR of breast cancer incidence and mortality, per 100 000 women: 40 y: 5-7 cases; 1.3-1.7 deaths 50 y: 2-3 cases; 0.7-0.9 deaths 80 y: 0.1-0.2 cases; <0.1 death LAR of breast cancer incidence and mortality in women undergoing annual screening mammography, per 100 000 women: Screening 40-80 y: 72-91 cases; 20-25 deaths Screening 50-80 y: 31-40 cases; 10-12 deaths
Yaffe and Mainprize, 2011 (92)	Modeling study	U.S.-based sources	40 to 74	Model based on digital mammography and radiation exposure estimates of 3.7 mGy per examination.	Estimated lifetime radiation-induced breast cancer cases and deaths	Number of radiation-induced breast cancer cases and deaths in 100 000 women: Annual screen 40-49 y: 59 cases; 7.6 deaths Annual 50-59 y: 27 cases; 3.1 deaths Biennial 50-59 y: 14 cases; 1.6 deaths Annual 40-59 y: 85 cases; 11 deaths Annual 40-49 y, biennial to 59 y: 73 cases; 9 deaths Annual 40-55 y, biennial to 74 y: 86 cases; 11 deaths

LAR = lifetime attributable risk.

Appendix Table 5. Studies of Harms of Breast Cancer Screening With Different Modalities

Author, Year (Reference)	Study Design	Population	Age, *y*	Study Period	Comparison (Number of Participants)	Outcome Measures	Results
Mammography with or without tomosynthesis							
Haas et al, 2013 (97)	Case series	United States; multisite hospital and outpatient centers	All ages	2011 to 2012	DM (7058) vs. DM plus tomosynthesis (6100)	Recall rate (%); adjusted odds of recall	Recall, DM vs. DM plus tomosynthesis, by age (relative change [95% CI]): All ages: 8.4% vs. 12.0%; −29.7% (−19.1% to −36.5%); $P < 0.01$ 40 to 49 y: 10.4% vs. 16.3%; −35.8% (−24.2% to −45.7%); $P < 0.01$ 50 to 59 y: 7.6% vs. 10.6%; −28.0% (−12.7% to −44.6%); $P < 0.01$ 60 to 69 y: 7.4% vs. 10.7%; −30.3% (−12.3% to −44.6%); $P = 0.01$ ≥70 y: 6.7% vs. 7.9%; −15.4% (NS) Adjusted recall OR (95% CI): 0.62 (0.55 to 0.70); $P < 0.0001$
Friedewald et al, 2014 (93)	Postintervention series	United States; multicenter	Mean: 57	2010 to 2012	DM (281 187) vs. DM plus tomosynthesis (173 663)	Recall and biopsy rates per 1000 women	Recall, DM vs. DM plus tomosynthesis (change [95% CI]): 107/1000 vs. 91/1000; −16.1 (−18.0 to −14.2); $P < 0.001$ Biopsy, DM vs. DM plus tomosynthesis (change [95% CI]): 18.1/1000 vs. 19.3/1000; 1.3 (0.4 to 2.1); $P = 0.004$
Rose et al, 2013 (94)	Case series	United States; multisite community-based breast center	>18	2011 to 2012	DM (18 202) vs. DM plus tomosynthesis (10 878)	Recall rate (%)	Recall, DM vs. DM plus tomosynthesis by age (relative change): All ages: 8.7% vs. 5.5%; −37.5%; NS <50 y: 10.3% vs. 6.5%; −37.2% 50–64 y: 7.6% vs. 5.1%; −32.9% >64 y: 7.9% vs. 4.2%; −46.6%
Ciatto et al, 2013 (95)	Postintervention series	Italy; population-based screening program (STORM)	≥48	2011 to June 2012	Biennial DM vs. DM plus tomosynthesis (total: 7292)	Recall rate (%)	Recall, DM vs. DM plus tomosynthesis: All ages: 141 (2%) vs. 73 (1%); $P < 0.0001$ <60 y: 89 (2.2%) vs. 41 (1.0%) >60 y: 52 (2%) vs. 32 (1%)
Skaane et al, 2013 (96)	Postintervention series	Oslo, Norway, screening program	50 to 69	2010 to 2011	Biennial DM vs. DM plus tomosynthesis (total: 12 631)	Recall rate per 1000 women	Recall, DM vs. DM plus tomosynthesis: 61.1/1000 vs. 53.1/1000 (−13%); RR, 0.85; $P < 0.001$
Mammography with or without CBE							
Chiarelli et al, 2009 (98)	Cohort	Canada	40 to 69	2002 to 2003	Biennial mammography (57 715) vs. CBE plus mammography (232 515)	Recall rate (%)	Recall, mammography vs. CBE with or without mammography: 6.5% vs. 8.7% (2.2% increase for CBE) or 55/10 000 additional FP results with CBE

CBE = clinical breast examination; DM = digital mammography; FP = false-positive; NS = not statistically significant; OR = odds ratio; RR = risk ratio; STORM = Screening with Tomosynthesis or Standard Mammography.