Evidence Synthesis
Number 124

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

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. HHSA-290-2012-00015-I, Task Order No. 2

Prepared by:
Pacific Northwest Evidence-based Practice Center
Oregon Health & Science University
Mail Code: BICC
3181 SW Sam Jackson Park Road
Portland, OR 97239

Investigators:
Heidi D. Nelson, MD, MPH
Amy Cantor, MD, MPH
Linda Humphrey, MD, MPH
Rochelle Fu, PhD
Miranda Pappas, MA
Monica Daeges, BA
Jessica Griffin, MS

AHRQ Publication No. 14-05201-EF-1
January 2016
This report is based on research conducted by the Pacific Northwest Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. HHSN-290-2012-00015-I, Task Order No. 2). The findings and conclusions in this document are those of the authors, who are responsible for its contents, and do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information (i.e., in the context of available resources and circumstances presented by individual patients).

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials that are clearly noted in the document. Further reproduction of those copyrighted materials is prohibited without the specific permission of copyright holders.

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Acknowledgements

The authors acknowledge Andrew Hamilton, MLS, MS, for conducting literature searches and Spencer Dandy, BS, for assistance with drafting this report at the Oregon Health & Science University; and Alison Conlin, MD, MPH, and Michael Neuman, MD, at the Providence Cancer Center, and Arpana Naik, MD, at the Oregon Health & Science University for medical expertise. The authors also thank the AHRQ Medical Officer, Jennifer Croswell, MD, MPH, and Tess Miller, DrPH, as well as current and former members of the U.S. Preventive Services Task Force who contributed to topic deliberations.

Data collection was supported by the National Cancer Institute-funded Breast Cancer Surveillance Consortium (BCSC) (HHSN-261-2011-00031-C, P01CA154292, and U54CA163303). Cancer data collection was supported in part by State public health departments and cancer registries in the United States (http://breastscreening.cancer.gov/work/acknowledgement.html). The authors thank the BCSC investigators and participating women, mammography facilities, and radiologists for the data they provided for this work, and especially Diana Miglioretti, PhD, Ellen O’Meara, PhD, and
Steven Balch of the BCSC Statistical Coordinating Center for providing information and analyses.

**Suggested Citation**

Structured Abstract

**Background:** In 2009, the U.S. Preventive Services Task Force (USPSTF) recommended biennial screening mammography for women age 50 to 74 years, and based decisions for earlier screening on individual patient context and values. Evidence was insufficient to recommend screening beyond age 75.

**Purpose:** To systematically update the 2009 USPSTF review on screening for breast cancer in average risk women age 40 years and older.

**Data Sources:** The Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews (through December 2014), Ovid MEDLINE (through December 2014), and reference lists were searched for relevant studies. Additional data were obtained from investigators of randomized trials and from the Breast Cancer Surveillance Consortium.

**Study Selection:** Randomized controlled trials and observational studies of breast cancer screening in asymptomatic women age 40 and older reporting breast cancer mortality, all-cause mortality, advanced breast cancer, treatment morbidity, and the harms of screening.

**Data Extraction:** One investigator abstracted data and a second investigator confirmed accuracy. Investigators independently dual-rated study quality and applicability using established criteria. Discrepancies were resolved through a consensus process.

**Data Synthesis:** A meta-analysis of screening trials with updated data from the Canadian (CNBSS-1 and CNBSS-2), Swedish Two-County Study, and Age trials indicated breast cancer mortality reductions for age 39 to 49 years (relative risk [RR] 0.88; 95% confidence interval [CI], 0.73 to 1.003; 9 trials; 4 deaths prevented/10,000 over 10 years); 50 to 59 years (RR 0.86 [95% CI, 0.68 to 0.97]; 7 trials; 8/10,000); 60 to 69 years (RR 0.67 [95% CI, 0.54 to 0.83]; 21/10,000); and 70 to 74 years (RR 0.80 [95% CI, 0.51 to 1.28]; 3 trials; 13/10,000). Risk reduction was 25 to 31 percent for women age 50 to 69 years across several observational studies, with similar reductions for women age 40 to 49 in two studies. Trials indicated no statistically significant reductions in all-cause mortality with screening. Risk for higher-stage breast cancer was reduced for age 50 years and older (RR 0.62 [95% CI, 0.46 to 0.83]; 3 trials), but not for age 39 to 49 years (RR 0.98 [95% CI, 0.74 to 1.37]; 4 trials). The majority of cases from screening were ductal carcinoma in situ and early stage, and screening resulted in more mastectomies (RR 1.20 [95% CI, 1.11 to 1.30]; 5 trials) and radiation (RR 1.32 [95% CI, 1.16 to 1.50]; 2 trials).

Younger women and those with risk factors had more false-positive results and recommendations for additional imaging and biopsies. Cumulative rates for false-positive mammography results over 10 years were 61 percent for annual and 42 percent for biennial screening; rates for biopsy were 7 to 9 percent for annual and 5 to 6 percent for biennial screening. Estimates of overdiagnosis ranged from 11 to 22 percent in trials; and 1 to 10 percent in observational studies. Some women with false-positive results or pain experienced distress and were less likely to return for their next mammogram. Tomosynthesis with mammography reduced recalls (16/1000), but increased biopsies (1.3/1000) and cancer detection (1.2/1,000).
The number of deaths due to radiation induced cancer from screening with digital mammography was estimated through modeling as between 2 to 11 per 100,000 depending on age at onset and screening intervals.

**Limitations:** Limited to English-language articles; the number, quality, and applicability of studies varied widely. Trials of mammography screening reflect imaging technologies and cancer treatment therapies that are not currently in use. Studies are lacking on screening effectiveness based on risk factors, intervals, and modalities; and on screening modalities relevant to women who are not high-risk.

**Conclusions:** Breast cancer mortality is reduced with mammography screening, although estimates are of borderline statistical significance, the magnitudes of effect are small for younger ages, and results vary depending on how cases were accrued in trials. Higher stage tumors are also reduced with screening for age 50 years and older. False-positive results are common in all age groups, and are higher for younger women and those with risk factors. Approximately 11 to 22 percent of cases may be overdiagnosed. Observational studies indicate that tomosynthesis with mammography reduces recalls, but increases biopsies and cancer detection. Mammography screening at any age is a tradeoff of a continuum of benefits and harms.
# Table of Contents

## Chapter 1. Introduction
- Purpose and Previous U.S. Preventive Services Task Force Recommendation .............................................. 1
- Condition Definition ............................................................................................................................................... 1
- Prevalence and Burden of Disease ..................................................................................................................... 2
- Etiology and Natural History............................................................................................................................... 2
- Risk Factors ......................................................................................................................................................... 2
- Rationale for Screening and Screening Strategies .............................................................................................. 3
- Interventions and Treatment ................................................................................................................................ 3
- Current Clinical Practice ....................................................................................................................................... 4
- Recommendations of Other Groups .................................................................................................................. 6

## Chapter 2. Methods
- Key Questions and Analytic Framework ............................................................................................................. 7
- Search Strategies .................................................................................................................................................. 8
- Study Selection .................................................................................................................................................... 9
- Data Abstraction and Quality Rating ................................................................................................................ 9
- Meta-Analysis of Mammography Screening Trials .......................................................................................... 10
- Analysis of BCSC Data ...................................................................................................................................... 10
- Data Synthesis ................................................................................................................................................... 11
- External Review ................................................................................................................................................ 11
- Response to Public Comment ........................................................................................................................... 11

## Chapter 3. Results
- Overview of Screening RCTs ............................................................................................................................ 13
- Key Question 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? ......................................................................................................................... 16
  - Summary ......................................................................................................................................................... 16
  - Evidence ....................................................................................................................................................... 17
- Key Question 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? ................................................................................................................................. 22
  - Summary ......................................................................................................................................................... 22
  - Evidence ....................................................................................................................................................... 23
- Key Question 3. How Does the Effectiveness of Routine Breast Cancer Screening in Reducing Breast Cancer–Specific and All-Cause Mortality Vary by Different Screening Modality? ...................................................................................................................... 25
  - Summary ......................................................................................................................................................... 25
- Key Question 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? ............................................................................................................................... 26
  - Summary ......................................................................................................................................................... 26
  - Evidence ....................................................................................................................................................... 26
- Key Question 5. What Are the Harms of Routine Mammography Screening, and How Do They Differ by Age, Risk Factor, and Screening Interval? ..................................................................................................................... 26
  - Summary ......................................................................................................................................................... 26
  - Evidence ....................................................................................................................................................... 28
Key Question 6. How Do the Harms of Routine Breast Cancer Screening Vary by Different Screening Modality? .................................................................38
Summary ..........................................................................................................................38
Evidence ............................................................................................................................38
Contextual Question 1. What Are the Rates of Specific Adverse Effects of Current Treatment Regimens for Invasive Breast Cancer and DCIS in the United States? ................................................39
Contextual Question 2. What Are the Absolute Incidence Rates of DCIS and Localized and Advanced Invasive Breast Cancer in Screened and Nonscreened Populations in the United States? .................................................................................40
Contextual Question 3. How Do Women Weight Harms and Benefits of Screening Mammography, and How Do They Use This Information in Their Decisions to Undergo Screening? .................................................................................40

Chapter 4. Discussion ........................................................................................................42
Summary of Review Findings ..........................................................................................42
Limitations .........................................................................................................................44
Emerging Issues and Next Steps ......................................................................................44
Relevance for Priority Populations ...................................................................................45
Future Research ..................................................................................................................45
Conclusions .........................................................................................................................45

References .........................................................................................................................47

Figures
Figure 1. Breast Cancer Screening Clinical Pathway
Figure 2. Clinical Pathway After Biopsy
Figure 3. Analytic Framework and Key Questions
Figure 4. Methods of Case Accrual in Trials
Figure 5. Meta-Analysis of Effects of Screening Trials on Breast Cancer Mortality With Longest Case Accrual Available
Figure 6. Meta-Analysis of Effects of Screening Trials on Breast Cancer Mortality With Short Case Accrual
Figure 7. Meta-Analysis of Effects of Screening Trials on All-Cause Mortality With Combined Ages
Figure 8. Meta-Analysis of Effects of Screening Trials on All-Cause Mortality Stratified by Age
Figure 9. Meta-Analysis of Effects of Screening Trials on Advanced Cancer Outcome Using a Low Threshold for Advanced Cancer
Figure 10. Meta-Analysis of Effects of Screening Trials on Advanced Cancer Outcome Using a High Threshold for Advanced Cancer

Tables
Table 1. Breast Cancer Staging System
Table 2. Treatment of DCIS and Invasive Breast Cancer by Stage
Table 3. Imaging Modalities for Breast Cancer Screening of Average-Risk Women
Table 4. Mammography Screening Trials
Table 5. Age-Specific Rates of Breast Cancer Mortality Reduction With Screening
Table 6. Biases and Limitations of Observational Studies of Mammography Screening
Table 7. Observational Studies of Screening and Mortality Not Included in Systematic Reviews
Table 8. Advanced Breast Cancer Outcomes Reported in Screening Trials
Table 9. Observational Studies of Advanced Cancer Outcomes With Mammography Screening
Table 10. Studies of Breast Cancer Treatment for Screened and Nonscreened Women
Table 11. Studies of Advanced Cancer Outcomes With Mammography Plus Tomosynthesis
Table 12. Age-Specific Rates of False-Positive and False-Negative Digital Mammography Results and Recommendations for Additional Imaging and Biopsies From a Single Screening Round in the BCSC
Table 13. Rates of False-Positive and False-Negative Digital Mammography Results and Recommendations for Additional Imaging and Biopsies Based on Time Since Last Mammography Examination
Table 14. Rates of False-Positive Results After Screening With Digital Mammography by Risk Factors
Table 15. Rates of False-Negative Results After Screening With Digital Mammography by Risk Factors
Table 16. Rates of Recommendations for Additional Imaging After Screening With Digital Mammography by Risk Factors
Table 17. Rates of Recommendations for Biopsy After Screening With Digital Mammography by Risk Factors
Table 18. Rates of False-Positive and False-Negative Digital Mammography Results and Recommendations for Additional Imaging and Biopsies by Different Breast Density Categories
Table 19. U.S. Studies of Cumulative False-Positive Mammography and Biopsy Results
Table 20. Studies of Overdiagnosis
Table 21. Overdiagnosis Estimates From Randomized, Controlled Trials Without Screening of Control Groups
Table 22. Systematic Reviews of Psychological Harms From False-Positive vs. Normal Screening Mammography Results
Table 23. Summary of Results of New Studies of Psychological Harms
Table 24. Models of Radiation Exposure, Breast Cancer Incidence, and Death
Table 25. Systematic Reviews of Pain With Mammography
Table 26. Studies of Harms of Screening With Different Modalities
Table 27. Rates of Specific Adverse Effects of Selected Treatments for Breast Cancer
Table 28. Summary of Evidence
Table 29. Summary of Results

Appendixes

Appendix A. Detailed Methods
Appendix A1. Search Strategies
Appendix A2. Inclusion and Exclusion Criteria per Key Question
Appendix A3. Literature Flow by Key Question
Appendix A4. List of Excluded Studies
Appendix A5. Quality Rating Criteria
Appendix A6. Details of the Breast Cancer Surveillance Consortium
Appendix A7. List of Peer Reviewers
Appendix A8. List of Studies Included in the Included Systematic Reviews
Appendix B. Evidence and Quality Tables
Appendix B1. Quality Ratings of Systematic Reviews
Appendix B2. Evidence Table of Results of New Studies of Psychological Harms
Appendix B3. Quality Ratings of Cohort Studies
Appendix B4. Quality Ratings for Case-Control Studies
Chapter 1. Introduction

Purpose and Previous U.S. Preventive Services Task Force Recommendation

This report will be used by the U.S. Preventive Services Task Force (USPSTF) to update their 2009 recommendation on screening for breast cancer. In 2009, the USPSTF recommended biennial screening mammography for women ages 50 to 74 years (B recommendation). They determined that the decision to start regular, biennial screening mammography before the age of 50 years should be an individual one and take patient context into account, including the patient's values regarding specific benefits and harms (C recommendation). The USPSTF concluded that evidence was insufficient to assess the additional benefits and harms of screening mammography in women age 75 years or older (I Statement).

The USPSTF also recommended against teaching breast self-examination (BSE) as a cancer screening strategy (D recommendation), and concluded that evidence was insufficient to assess the additional benefits and harms of clinical breast examination (CBE) beyond screening mammography in women age 40 years or older (I Statement). The USPSTF concluded that the evidence was insufficient to assess the additional benefits and harms of either digital mammography or magnetic resonance imaging (MRI) instead of film mammography as screening modalities for breast cancer (I Statement).

This report updates evidence on the effectiveness of mammography in decreasing breast cancer mortality, all-cause mortality, and advanced breast cancer among women who are not at high risk for breast cancer; harms of screening; and how effectiveness and harms vary by age, risk factors, screening intervals, and screening modalities. This report includes studies relevant to current medical practice in the United States and highlights gaps as well as strengths in evidence. Additional reviews and analyses for the USPSTF are provided in separate reports including systematic reviews of the performance characteristics of screening methods and the accuracy of breast density determination and use of supplemental screening technologies, and a model of radiation exposure.

Condition Definition

Breast cancer is a proliferation of malignant cells that arises in the breast tissue, specifically in the terminal ductal-lobular unit, and represents a continuum of disease ranging from noninvasive to invasive carcinoma. Noninvasive carcinoma, or an in situ lesion, does not invade the surrounding stroma and does not metastasize. Noninvasive lesions are confined to either the duct (ductal carcinoma in situ [DCIS]), or to the lobule (lobular carcinoma in situ [LCIS], now categorized as lobular intraepithelial neoplasia [LIN]). LCIS is considered a marker for increased risk of invasive ductal or lobular breast cancer, while some forms of DCIS are considered precursor lesions for invasive ductal carcinoma. DCIS is heterogeneous and has varying clinical behavior and pathologic characteristics.
Unlike noninvasive lesions, invasive breast cancer invades the basement membrane into the adjacent stroma, and therefore, has metastatic potential. The most common sites of metastasis include adjacent lymph nodes, lung, liver, and bone. Approximately 70 to 80 percent of invasive breast cancer cases are invasive or infiltrating ductal carcinoma and approximately 10 percent are invasive lobular carcinoma. Other less common histologic subtypes of invasive breast cancer include apocrine, medullary, metaplastic, mucinous, papillary, and tubular.

**Prevalence and Burden of Disease**

Breast cancer is the second most common cancer in women in the United States after non-melanoma skin cancer, and is the second leading cause of cancer death after lung cancer. In 2015, an estimated 231,840 women in the United States will be diagnosed with breast cancer and 40,290 will die, representing 14 percent of all new cancer cases and 6.8 percent of all cancer deaths. Incidence rates have been stable over the last 10 years and death rates have been falling approximately 1.9 percent each year between 2002 and 2011. According to lifetime risk estimates for the general population, 12.3 percent of women will develop breast cancer during their lives, and 2.8 percent will die from the disease. The overall 5-year relative survival rate for breast cancer in 2006 was 90.6 percent, and an estimated 2,899,726 women were living with breast cancer in the United States in 2011.

**Etiology and Natural History**

Current research on the etiology of breast cancer focuses on clarifying the role of both inherited and acquired mutations in oncogenes and tumor suppressor genes and the consequences these mutations may have on the cell cycle, as well as investigating various prognostic biological markers. The contribution of external influences, such as environmental exposures, have on regulatory genes is unclear. Currently, no single environmental or dietary exposure has been found to cause a specific genetic mutation that causes breast cancer. Exposure to both endogenous and exogenous estrogen is important in tumorigenesis and growth. Other potential causes of breast cancer include inflammation and virally mediated carcinogenesis.

Whether DCIS is a precursor lesion or a marker of risk is uncertain. With the widespread use of screening mammography in the United States, nearly 90 percent of DCIS cases are now diagnosed only on imaging studies, most commonly by the presence of microcalcifications. These represent approximately 23 percent of all breast cancer cases. Although DCIS is the most common type of noninvasive breast cancer, its natural history is poorly understood. Older studies of palpable DCIS lesions indicated that 14 to 53 percent of untreated DCIS progressed to invasive cancer over 8 to 22 years. The rate of progression of mammography detected DCIS is not known. Characteristics associated with subsequent invasive breast cancer include young age, black race, indication for biopsy, tumor characteristics such as high nuclear grade, comedo-type necrosis, tumor size; and high breast density.
Risk Factors

Although many risk factors have been associated with breast cancer in epidemiologic studies, most relationships are weak or inconsistent.\(^{17}\) Most women who develop breast cancer have no identifiable risk factors beyond sex and age. However, a small number of clinically significant risk factors are associated with high risks for breast cancer and can be used to identify women who may be eligible for screening outside routine screening recommendations. These include women with \textit{BRCA1} or \textit{BRCA2} mutations and their untested first-degree relatives,\(^{18}\) and other hereditary genetic syndromes associated with more than a 15 percent lifetime risk, including Li-Fraumeni syndrome, Cowden syndrome, or hereditary diffuse gastric cancer.\(^{19}\) Previously diagnosed high-risk breast lesions, including LCIS, atypical ductal hyperplasia (ADH), atypical lobular hyperplasia (ALH), flat epithelial atypia, papillary atypia, and apocrine atypia significantly increase risk for breast cancer.\(^{20}\) Estimated 10-year breast cancer risks associated with breast lesions include 17.3 percent with ADH, 20.7 percent with ALH, 23.7 percent with LCIS, and 26.0 percent with severe ADH.\(^{20}\) Also, women with a history of high-dose radiation therapy to the chest between the ages of 10 to 30 years, such as for treatment of Hodgkin lymphoma, are also considered at high risk.\(^{19}\)

Family history of breast cancer, particularly among first-degree relatives, is also an important risk factor. Approximately 5 to 10 percent of women with breast cancer have a mother or sister with breast cancer, and up to 20 percent have either a first-degree or a second-degree relative with breast cancer.\(^{21-25}\) The degree of risk associated with family history varies according to familial patterns of disease. Estimates of lifetime risk of breast cancer determined by kindred analysis of over 15 or 20 percent are considered high.

Additional factors that increase risk to lower degrees than described above include older age; current use of menopausal hormone therapy using combined estrogen and progestin regimens;\(^{26}\) current use of oral contraceptives;\(^{17}\) nulliparity;\(^{17}\) high body mass index (BMI) for postmenopausal women only;\(^{27}\) and higher breast density.\(^{28}\) Breast density is a radiographic measure of breast tissue that is associated with increased risk for breast cancer and reduced mammography sensitivity. Breast density is currently described by four categories: almost entirely fat, scattered fibroglandular densities, heterogeneously dense, and extremely dense.\(^{29}\) Approximately 40 percent of women have heterogeneously dense breasts and 10 percent have extremely dense breasts. Increased breast density is more common among younger women.\(^{30}\) Compared with women with scattered fibroglandular densities, hazard ratios for breast cancer are 1.6 for premenopausal women with heterogeneously dense breasts and 2.0 for those with extremely dense breasts.\(^{28}\)

Empiric models that incorporate several of these risk factors have been developed to predict breast cancer risk for individual women.\(^{31}\) All of the models include age and number of first-degree relatives with breast cancer into their calculations, but vary in their complexity. Studies of their diagnostic accuracy indicate that the models are poor predictors of an individual’s risk.\(^{31}\) It remains unclear how to apply these models to selecting candidates for breast cancer screening.
Rationale for Screening and Screening Strategies

Breast cancer has a known asymptomatic phase that can be identified with mammography and could be more effectively treated in early stages than when clinical signs and symptoms present. While screening may not reduce mortality for some aggressive cancer types, and has less impact on slowly progressive types, survival may be improved for other types of cancer when they are identified at localized stages.

Interventions and Treatment

Current treatment for breast cancer in the United States involves a combination of therapies including surgery, radiation, hormonal therapy, and chemotherapy based on stage (0 to IV) and status of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). Clinical staging using the American Joint Committee on Cancer (AJCC) TNM system guides treatment and informs prognosis (Table 1). In this system, stage grouping is based on tumor size (T), lymph node involvement (L), and presence of metastasis (M). Main categories are expressed as DCIS (stage 0), localized (Stage I, IIA, IIB, or T3, N1, M0), locally advanced or regional (Stage III), and metastatic disease (Stage IV). Survival varies by stage, and the 5-year relative survival rates for breast cancer in the United States are 99 percent with localized, 84 percent with regional, and 23 percent with metastatic disease.

Treatment regimens are highly individualized according to each patient’s clinical status and preferences (Table 2). In addition, many patients are recruited to clinical trials of new regimens. Surgical therapies for DCIS and localized and regional invasive cancer include lumpectomy or total mastectomy with or without reconstruction. Surgery also involves sentinel lymph node biopsy for selected cases of DCIS, axillary node staging for localized disease, and axillary node dissection for regional disease. Surgical therapy is performed in only selected cases of metastatic disease. Radiation therapy generally follows surgery. Whole breast radiation may be added to lumpectomy for DCIS, localized, and regional disease. Radiation to the lymph nodes and chest wall, if involved, may be indicated for localized and regional disease.

Endocrine treatment is recommended for ER-positive patients at all stages. Usual regimens include 5 years of tamoxifen for DCIS, 10 years of extended adjuvant hormonal therapy (tamoxifen with or without aromatase inhibitors) for localized and regional disease, and additional regimens for metastatic disease. Premenopausal ER-positive patients with metastatic disease may also consider ovarian ablative or suppression.

Systemic neo-adjuvant/adjuvant chemotherapy for invasive cancer is determined by ER, PR, and HER2 status, and predictive tests for chemotherapy benefit. Chemotherapy is given after surgery for localized disease. Patients with regional disease generally receive chemotherapy before or after breast surgery and incorporate one year of trastuzumab (Herceptin). Chemotherapy for metastatic disease involves more complicated regimens depending on receptor status, tumor biology, and initial responses.
Current Clinical Practice

Mammography screening in the United States is generally opportunistic, unlike the many screening programs organized as public health services in other countries. Mammography is provided by radiology units of hospitals and outpatient facilities as well as by stand-alone imaging centers. Services range from imaging alone to comprehensive services that may be integrated within breast centers. As such, there is considerable variation in current clinical practice depending on the patient population, provider practice, community, and institutional policy, although national accreditation and professional groups define practice standards and quality benchmarks to assure consistency of care.

While there is general consensus that mammography screening is beneficial for many women, conflicting screening recommendations have led to practice variability. Issues lacking consensus include the optimal ages to begin and end routine screening; optimal screening intervals; defining and balancing the benefits of screening with potential harms; appropriate use of various imaging modalities including supplemental technologies; values and preferences of women regarding screening; and how all of these considerations vary depending on a woman’s risk for breast cancer.

Despite variation in clinical practices and guidelines, rates of screening mammography in the United States are generally high and have remained relatively stable for the past decade. Data from the HEDIS® Health Plan measure set indicate that mammography screening between 2009 and 2011 was performed by 71 percent of eligible women covered by commercial plans, 69 percent for Medicare plans, and 51 percent for Medicaid plans. The Patient Protection and Affordable Care Act mandates insurance coverage for annual screening mammography beginning at age 40 years with no co-pay or deductible charges. However, this coverage applies to only the annual screening mammogram, and subsequent related services are not similarly covered.

Breast cancer screening for women without risk factors indicating high risk is conducted using periodic mammography (Figure 1). Digital mammography has generally replaced film in the United States, and newer technologies, such as digital tomosynthesis, are rapidly disseminating. Imaging modalities are further described in Table 3. In general, approximately 90 percent of women in a screening round have normal mammography results and are advised to return in 1 or 2 years, while 10 percent are recalled for additional imaging to visualize areas of concern identified on the screening mammogram. Additional imaging may involve special mammographic views, ultrasound, MRI, or tomosynthesis. Approximately 10 percent of women having additional imaging are identified with suspicious breast lesions requiring biopsies.

Additional imaging after screening mammography has traditionally been reserved to further visualize incompletely evaluated breast lesions. However, in response to public concerns about breast density, 21 states have passed breast density notification legislation requiring that reports of patients’ breast density be provided to them with their mammography results. Most laws encourage patients to have a discussion of additional screening options with their primary physicians and some mandate insurance coverage for supplemental imaging, including screening MRI, ultrasound, and tomosynthesis. Descriptive studies of supplemental imaging for patients...
with dense breasts, in addition to other risk factors in some studies, suggest increased rates of
cancer detection, but also increased false-positive results with MRI and ultrasound.\textsuperscript{46-50} Randomized trials of the effectiveness of supplemental imaging have not been reported.\textsuperscript{51}

Screening MRI is recommended for certain high-risk groups, including women with \textit{BRCA1} or \textit{BRCA2} mutations and their untested first-degree relatives, women with greater than 20 percent lifetime risk of developing breast cancer as defined by risk prediction models, and women who have received high-dose radiation therapy to the chest between the ages of 10 and 30 years.\textsuperscript{52} Use of MRI for screening women who are not at high risk for breast cancer is not recommended,\textsuperscript{52} and experts suggest that MRI should not be performed in settings where the capacity for MR-guided biopsy does not exist. Currently, there are no studies investigating MRI use in women who are not at high risk, and none showing decreased mortality with MRI screening for women at any risk level.

If tissue sampling is recommended, a biopsy is performed (Figure 2). The type of biopsy is based on the characteristics of the lesion as well as patient and physician preferences. Current biopsy techniques include fine-needle aspiration (FNA), stereotactic core biopsy (for nonpalpable, mammographic lesions), ultrasound-guided or MRI-guided core biopsy, non-image-guided core biopsy (for palpable lesions), incisional biopsy, or excisional biopsy. These techniques vary in the level of invasiveness and amount of tissue acquired, impacting their yield and patient experience. Although more invasive than FNA, core biopsies, as well as incisional and excisional biopsies, offer the pathologist a sample with intact cellular architecture, and thereby allow additional pathologic examination of the breast tissue. Testing includes examination of cellular receptors (e.g., ER/PR, HER2/neu receptor), as well as identification of tumor type and grade.\textsuperscript{53,54} Ultrasound of the ipsilateral axilla has become a common practice when malignancy is suspected on imaging, and can help guide FNA or core biopsy of abnormal axillary lymph nodes. This additional information contributes to appropriate treatment planning for a patient who is newly diagnosed with breast cancer, and often allows for definitive surgery to be completed with a single-stage procedure.\textsuperscript{55}

**Recommendations of Other Groups**

The American Cancer Society recommends yearly mammograms starting at age 40 and continuing for as long as the woman is in good health. BSEs are optional for women beginning in their 20s, while CBEs are recommended about every 3 years for women in their 20s and 30s.\textsuperscript{56}

The American Academy of Family Physicians (AAFP) recommends the decision to conduct screening mammography prior to age 50 should be individualized and take into consideration the patient’s context and risk factors. For women between ages 50 and 74, the AAFP recommends biennial screening in addition to recommending against clinicians teaching women BSE.\textsuperscript{57}

The American Congress of Obstetrics and Gynecologists (ACOG) recommends that mammography screening be offered annually to women beginning at age 40. ACOG recommends annual CBE for women ages 40 and older, and every 1 to 3 years for women ages 20 to 39. ACOG also endorses educating women ages 20 and older regarding breast self-
The American College of Radiology (ACR) recommends annual screening mammography for asymptomatic women 40 years of age and older. The decision as to when to stop routine mammography screening should be made on an individual basis by each woman and her physician based on a woman’s overall health.

The National Comprehensive Cancer Network (NCCN) recommends annual screening mammography, clinical breast exam, and breast awareness for asymptomatic, average risk women age 40 years and older.
Chapter 2. Methods

Key Questions and Analytic Framework

Using the methods developed by the USPSTF,61,62 the USPSTF and the Agency for Healthcare Research and Quality (AHRQ) determined the scope and key questions for this review. Investigators created an analytic framework outlining the key questions and included patient populations, interventions, and outcomes (Figure 3).

The target population for the USPSTF recommendation served as the focus of the systematic review. This population includes women age 40 years and older and excludes women with 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. Women at high-risk are those with risk factors known to increase their risks of breast cancer to levels that make them eligible for screening or followup services outside of recommendations for women without these risk factors. Women at high-risk include those with pre-existing breast cancer; \( BRCA1 \) or \( BRCA2 \) mutations and their untested first-degree relatives18 and other hereditary genetic syndromes associated with more than a 15 percent lifetime risk of developing breast cancer (including Li-Fraumeni syndrome, Cowden syndrome, or hereditary diffuse gastric cancer);19 previously diagnosed high-risk breast lesions (DCIS, LCIS, ADH, ALH); and high-dose radiation therapy to the chest between the ages of 10 and 30. Women with lower risks for breast cancer are generally eligible for routine screening and are relevant to the USPSTF’s recommendations.

Key Questions

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 factors, and screening intervals?
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 factors, and screening intervals?
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?
5. What are the harms of routine mammography screening, and how do they differ by age, risk factors, and screening intervals?
6. How do the harms of routine breast cancer screening vary by different screening modality?

Risk factors considered in this review are common among women who are not at high-risk for breast cancer as defined above. These include family history of breast cancer (not including genetic syndromes described above), breast density, race/ethnicity, menopausal status, current...
Outcomes related to benefits included in this review are reduced breast cancer mortality, all-cause mortality, advanced breast cancer, and morbidity related to breast cancer treatment. Other outcomes, such as increased breast cancer awareness and peace of mind with screening, are not included. Treatment-related morbidity includes physical adverse effects of treatment, quality of life measures, and other measures of impairment. Screening modalities include mammography (digital, tomosynthesis), MRI, ultrasound, and CBE (alone or in combination). Only breast imaging technologies approved for screening by the U.S. Food and Drug Administration are included in this review, consistent with the scope of the USPSTF.

Harms include false-positive and false-negative mammography results, false reassurance, anxiety and worry, overdiagnosis and resulting overtreatment, and radiation exposure. Overdiagnosis refers to women receiving a diagnosis of DCIS or invasive breast cancer who had abnormal lesions that were unlikely to become clinically evident during their lifetimes in the absence of screening. Overdiagnosis may have more effect on women with shorter life expectancies because of age or comorbid conditions.

**Contextual Questions**

Three contextual questions were also requested by the USPSTF to provide additional background information. Contextual questions are not reviewed using systematic review methodology but are addressed using the strongest, most relevant evidence. These include the following.

1. What are the rates of specific adverse effects of current treatment regimens for invasive breast cancer and DCIS in the United States?
2. What are the absolute incidence rates of DCIS and localized and advanced invasive breast cancer in screened and nonscreened populations in the United States?
3. How do women weigh the harms and benefits of screening mammography and how do they use this information in their decisions to undergo screening?

**Search Strategies**

In conjunction with the systematic review investigators, a research librarian searched the Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and Ovid MEDLINE (all searches through December 2014) for relevant studies and systematic reviews. Reference lists of articles were also reviewed. Search dates varied because some key questions (Key Questions 1, 3, 5, 6) were included in the 2009 systematic review and required only updates of studies published since the previous search in 2008. Other key questions were not addressed by the previous review and required searches that covered longer time periods (Key Questions 2 and 4, and cohort studies for Key Questions 1 and 3). These searches extended to 1996 because this corresponds to the last time the USPSTF evaluated similar data, and represents a period when practice was shifting to digital mammography in the United States. The contextual questions have a shorter time period for searches because they require the most
In addition, unpublished data from the Breast Cancer Surveillance Consortium (BCSC) on screening with digital mammography were evaluated. The BCSC is a collaborative network of mammography registries with linkages to pathology databases and tumor registries across the United States supported by the National Cancer Institute (NCI). These data draw from community samples that are representative of the larger, national population and may be more applicable to current practice in the United States than other published sources.

Also, unpublished updated data from the Canadian and Swedish Two-County randomized controlled trials (RCTs) were obtained from the trial investigators.

### Study Selection

At least two reviewers independently evaluated each study to determine inclusion eligibility. Studies were selected on the basis of prespecified inclusion and exclusion criteria developed for each key question (Appendix A2). The selection of studies is summarized in a flow diagram (Appendix A3). Appendix A4 lists excluded studies with reasons for exclusion.

Studies of women at high-risk for breast cancer as defined above or with previously diagnosed breast cancer were not included. 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, use of technologies and therapies in current practice, and other factors. Also, studies of higher-quality and those with designs ranked higher in the study design-based hierarchy of evidence, such as RCTs over observational studies, were emphasized because they are less susceptible to bias.

To determine the effectiveness of screening, RCTs, observational studies of screening cohorts, and systematic reviews of screening with mammography (film, digital, tomosynthesis) and other modalities (MRI, ultrasound, CBE alone or in combination) were included. Valid comparisons evaluated outcomes of groups of women exposed to screening versus nonscreening, not comparisons of detection methods that do not capture a woman’s longitudinal screening experience (e.g., rates of screen-detected vs. nonscreen-detected cancer).

Outcomes included breast cancer specific and all-cause mortality (Key Questions 1 and 3) and advanced breast cancer and treatment-related morbidity (Key Questions 2 and 4). While advanced breast cancer is classified as metastatic disease (Stage IV) by the AJCC TNM system, most screening studies defined advanced breast cancer at much lower thresholds, including Stage IIA or higher, lymph node positive disease, or tumor size of 20 mm or larger. Studies providing outcomes specific to age, risk factors, screening intervals, and modalities were preferred over studies providing general outcomes, when available. Risk factors conferring a moderate, as opposed to high, level of risk were included as listed previously.

The harms of screening were determined from several study designs and data sources. For mammography, searches focused on recently published systematic reviews and meta-analyses of
radiation exposure, pain during procedures, patient anxiety and other psychological responses, consequences of false-positive and false-negative mammography results, and overdiagnosis. Specific searches for primary studies published more recently than the included systematic reviews and meta-analyses were also conducted.

Performance characteristics of screening methods (e.g., sensitivity, specificity, positive predictive value); accuracy of breast density determination; use of supplemental screening technologies; and a new model of radiation exposure are presented in separate reports. Studies of cost-effectiveness of screening were not addressed in this update.

Data Abstraction and Quality Rating

Details of the study design, patient population, setting, screening method, interventions, analysis, followup, and results were abstracted by one investigator and confirmed by a second. Two investigators independently applied criteria developed by the USPSTF\textsuperscript{61,62} 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 (Appendix A5). USPSTF criteria to rate other study designs included in this review are not available. Discrepancies were resolved through consensus. Only data from RCTs rated fair- or good-quality were included in the meta-analyses.

Meta-Analysis of Mammography Screening Trials

Several meta-analyses were conducted to determine more precise summary estimates for the effectiveness of breast cancer screening when adequate data were reported by trials. Clinical and methodological diversity and statistical heterogeneity were considered to determine the appropriateness of meta-analysis. All outcomes (breast cancer mortality, all-cause mortality, and advanced cancer occurrence) were binary. A random-effects model was used 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.\textsuperscript{67} The presence of statistical heterogeneity among the studies was assessed by using the standard Cochran’s chi-square test, and the magnitude of heterogeneity by using the $I^2$ statistic.\textsuperscript{68}

To account for clinical heterogeneity and obtain clinically meaningful estimates, the analyses were stratified by age group (39 to 49 years, 50 to 59 years, 60 to 69 years, 70 to 74 years, or $\geq$75 years), whenever possible. Investigators of two recently published updates of trials provided additional age-stratified data for the meta-analysis.\textsuperscript{69,70} Two definitions were used to evaluate advanced breast cancer outcomes (stage and tumor size).

For breast cancer mortality, two methods of including cases in estimates were used because each offers advantages and disadvantages, and may provide additional insights to the interpretation of results. The long case accrual method counts all of the breast cancer cases contributing to breast cancer deaths. In this method, the case accrual time is equivalent to or close to the followup 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. These methods are further described in the results section.

To facilitate the interpretation of the combined RR for breast cancer mortality, the absolute rate reduction for 100,000 women-years of followup (i.e., 10,000 women followed for 10 years) was calculated for each age group based on the combined RR and the combined cancer rate of the control group. The combined cancer rate of the control group was obtained using a random effects Poisson model for each age group using data from the trials. All analyses were performed using Stata/IC 13.1 (StataCorp, College Station, TX).

Analysis of BCSC Data

Background information and additional details about methods of the BCSC are described in Appendix A6. Data were obtained from the BCSC Statistical Coordinating Center for 405,191 women ages 40 to 89 years who had routine screening with digital mammography during 2003 to 2011 at participating facilities at six BCSC breast imaging registries. Results were stratified by age in decades to determine age-specific outcomes. Routine screening required at least one mammography examination within the previous 2 years (defined as 30 months). For women with several mammography examinations during this time period, one result was randomly selected to be included in the calculations. These data comprise a defined subset of BCSC data intended to represent the experience of a cohort of regularly screened women without histories of breast cancer or current breast symptoms.

Screening mammography examinations were those designated as such by the radiologist or radiology technologist performed more than 9 months after a previous imaging examination in women without histories of breast cancer, breast augmentation, or mastectomies. This approach eliminated the possibility that a woman’s first mammogram was included because first mammograms are more likely to be read as false positives. Unilateral exams were also excluded. Mammography information included Breast Imaging Reporting and Data System (BI-RADS) breast density, assessment scores, and recommendations for further workup. In addition, prior to each mammography examination, women completed questionnaires that included demographic and medical history information, including previous mammography information.

Data include the numbers of positive and negative mammography results and, of these, the numbers of normal screening and false-negative results based on followup data within 1 year of mammography screening and before the next screening examination. Positive versus negative initial and final results were defined according to standardized terminology and assessments of the American College of Radiology BI-RADS 4th edition atlas and BCSC standard definitions. Each screening mammography examination was given an initial BI-RADS assessment based on the screening views only. Positive initial results included four assessment categories: needs additional imaging evaluation (category 0), probably benign (category 3) with a recommendation for immediate work-up (these were treated as a category 0 based on the recommendation), suspicious abnormality (category 4), or highly suggestive of malignancy (category 5). Negative results included assessments of negative (category 1) or benign findings (category 2), or category 3 without a recommendation for immediate work-up.
For women who had positive screening mammography results, data were evaluated on the number of women receiving a recommendation for additional imaging, the number receiving a recommendation for biopsy, and diagnoses including invasive breast cancer, DCIS, and no cancer. Recommendation for biopsy was defined as a positive final result after all imaging including work-up of an abnormal screening examination. Positive final results included BI-RADS assessments of 4 or 5 or 0 with a recommendation for biopsy. Negative final results included an assessment of 1, 2, or 3 or 0 with a recommendation for normal or short-interval followup or clinical exam.

From these data, age-specific rates (numbers per 1,000 women per screening round) of invasive breast cancer, DCIS, false-positive and false-negative mammography results, recommendations for additional imaging, and recommendations for biopsies were calculated.

Age groups were further divided into sub-categories to determine whether outcomes differed by time since last mammography screening or risk factors. Two measures of time since last mammography screening were evaluated to represent broad and narrow estimates of one versus two years (9 to 18 versus 19 to 30 months; 11 to 14 versus 23 to 26 months).

Risk factors included those mostly commonly associated with breast cancer. These included first-degree relatives with breast cancer (none, ≥1); breast density (almost entirely fat, scattered fibroglandular densities, heterogeneously dense, extremely dense); benign breast biopsy (none, previous); race/ethnicity (white, black, Asian, Hispanic, other); menopausal status (pre, peri, postmenopausal); menopausal hormone therapy use (none, combination [estrogen with progestin], estrogen only); oral contraceptive use (no current use, current use), and body mass index (BMI) (<25, 25 to <30, ≥30 kg/m²). Since the BCSC data do not include information on types of menopausal hormone therapy, the analysis assumes that a woman with a uterus uses combination therapy, while a woman without a uterus uses estrogen-only therapy. The main analysis analyzed three categories of breast density, combining almost entirely fat and scattered fibroglandular densities into one group. As a sensitivity analysis, density was analyzed in three additional ways: (1) three categories, combining heterogeneously dense and extremely dense into one group; (2) four separate BI-RADS categories; and (3) two categories that combine almost entirely fat and scattered fibroglandular densities into one group and heterogeneously dense and extremely dense into another group.

From these data, age-specific rates (numbers per 1,000 women per screening round) of false-positive and false-negative mammography results, recommendations for additional imaging, and recommendations for biopsies were calculated and comparisons by age, time since last mammography screening, and risk factors were determined. To account for correlation among mammograms interpreted at the same radiology facility, robust standard errors from logistic regression were estimated using generalized estimating equations with an independence working correlation matrix. Differences between groups were assessed by 2-sided P-values with 95% confidence intervals (CIs).
Data Synthesis

The aggregate internal validity (quality) of the body of evidence for each key question was assessed ("good," "fair," "poor") using methods developed by the USPSTF, based on the number, quality and size of studies, consistency of results between studies, and directness of evidence.61,62

External Review

The draft report was reviewed by content experts from multiple disciplines, USPSTF members, AHRQ Project Officers, and collaborative partners and revised prior to finalization (Appendix A7).

Response to Public Comment

A draft version of the evidence report was posted for public comment on the USPSTF website from April 18 to May 18, 2015. Comments from 13 contributors were directly relevant to the systematic review, while comments from other contributors were outside its scope, or concerned the recommendation statement, CISNET model, or other evidence reports. Most comments addressed four major issues detailed below, while additional comments suggested adding studies that were either already included or were previously considered, but did not meet prespecified inclusion criteria; or correcting minor errors that have since been corrected.

Inclusion of Observational Studies

The evidence report includes nearly 200 observational studies of breast cancer screening including 83 studies of benefits. Results of studies of the effectiveness of screening in reducing breast cancer mortality are reported for both observational studies and RCTs. For women age 50 to 69 years, the trials indicated statistically significant reductions in breast cancer mortality ranging from 0.78 to 0.81 depending on whether short or long case accrual methods were used. Observational studies indicated reductions of 0.69 to 0.75. For women age 40 to 49, few data from observational studies were available because most European countries collecting these data do not screen younger women. For women age 70 and older, data from both RCTs and observational studies were not available.

RCTs are the least biased study design for determining efficacy/effectiveness, and provide a stronger body of evidence than observational studies. When RCTs and studies of other designs have similar results, such as breast cancer mortality reduction for women age 50 to 69 years, the body of evidence is stronger. Observational studies are subject to important biases that limit their use in determining effectiveness. Most importantly, they lack comparability of comparison groups that is only attainable through randomization. Many observational studies that compare characteristics of breast cancer diagnoses between screened and unscreened women provide comparisons between screen-detected and nonscreen-detected cases. This approach categorizes all cancer cases identified outside of a screening mammogram as nonscreen-detected,
even though a woman may have had prior screening mammography.

In RCTs, intension-to-screen analysis is essential to determining efficacy/effectiveness, and is comparable to intension-to-treat analysis for drug trials. Data from trials not using intension-to-screen analysis, or from observational studies, provide outcomes for women who self-select screening. While outcomes from women who self-select screening may be useful for planning health care services, they do not provide valid measures of efficacy/effectiveness. This is a major difference between how evidence-based guideline groups and some professional societies interpret the research literature.

**Inclusion of RCTs**

The evidence review describes the RCTs of screening and their limitations in detail. No trials met criteria for good quality (all RCTs in the meta-analysis were fair-quality; a poor-quality trial was excluded).

The meta-analyses of RCTs for breast cancer mortality outcomes use two methods in order to more precisely explain the results of the trials and provide a range of outcomes. There are advantages and disadvantages to these methods, and these are described in the evidence review. While both methods have been used for individual trials and for some of the Swedish trials collectively, no other systematic reviews have taken this rigorous approach across all trials. Results of some of the trials appear in both estimates because the trial investigators only published short case accrual results. Rather than eliminate trials from the meta-analysis, the “longest followup available” from each trial was included and those based on short case accrual are clearly indicated (this was also the approach in the 2009 meta-analysis).

Regarding outcomes related to advanced breast cancer, most of the diagnostic outcomes of the trials were based on early stages of disease (Stage IIA or localized), not advanced. To address the key question about prevention of advanced disease, the meta-analysis used the most advanced disease categories available from the trials. These results indicated reduced risk with screening for women age 50 and older. The connection between being diagnosed with advanced breast cancer and dying is not a key question of this evidence review. This link is acknowledged in the analytic framework.

**Screening Intervals**

None of the RCTs were designed to evaluate screening intervals. The observational data, including studies from the BCSC, are based on women who self-select screening and adhere, or not, to specific periods of time between screening. Comparisons between women who electively screen annually versus biennially are inherently biased because these women differ in many ways. Estimates from BCSC data are approximations that reflect opportunistic screening in a fluctuating population of women whose information was collected by the participating registries. The BCSC data need to be interpreted with these limitations in mind.
Harms

The studies on overdiagnosis are described in detail in the evidence review and the general conclusion is that they are too methodologically heterogeneous to provide reliable estimates. Until a consensus definition with common metrics is determined, these estimates are uncertain.
Chapter 3. Results

Overview of the RCTs of Screening

Eight main RCTs of mammography screening provide outcomes that address several key questions for this review. Trials involving over 600,000 women have been conducted in the United States, Canada, United Kingdom, and Sweden. These include the Health Insurance Plan of Greater New York (HIP) trial,74 Canadian Breast Cancer Screening Study 1 (CNBSS-1),75,76 Canadian Breast Cancer Screening Study 2 (CNBSS-2),77,78 United Kingdom Age trial,79 and four from Sweden, including the Stockholm trial,80 Malmö Mammographic Screening Trial (referred to separately as MMST I and MMST II),81 Gothenburg trial,82 and Swedish Two-County Study (referred to separately as Östergötland and Kopparberg).83 All of these trials met criteria for fair quality and were included in this report. An additional trial, the Edinburgh trial,84,85 was not included in this review because of its inadequate randomization, introducing high risk of bias and limiting any inferences.

Updates of three trials provided new data for this report,69,70,86 although only the Canadian and Age trials provided published results that were stratified by age groups.69,86 Age-stratified results for the Swedish Two-County Study were provided by the trial investigators (Dr. Lászlo Tabár personal communication).

Trials varied in their recruitment of participants, screening protocols, control groups, and sizes (Table 4). The HIP trial used direct-exposure film mammography, while all of the other trials used screen-film mammography, and none evaluated digital mammography or tomosynthesis. Five trials examined the effectiveness of screening among women between the ages of 40 and 74 years;74,80-84 two trials enrolled only women in their 40s;75,79 and one enrolled only women in their 50s.77 The four trials from Sweden and the Age trial from the United Kingdom evaluated mammography alone, and the other trials evaluated the combination of mammography and CBE. Overviews of the Swedish trials providing outcome data have also been published.87,88 The overviews addressed several important study limitations of the Swedish trials including reassessing causes of death in the Swedish Two-County Study with a blinded independent end point committee.

Five trials were randomized at the individual participant level (CNBSS-1, CNBSS-2, HIP, Age, Stockholm, and Malmö); one trial used individual (82%) and cluster (18%) randomization (Gothenburg); and two trials used cluster randomization by community (Swedish Two-County). Breast cancer mortality was the main outcome measure, and all trials evaluated differences between the screening and control groups on an intention-to-screen basis. Seven studies randomized women to an invitation to screening or control group receiving “usual care” at the time the study was conducted. Usual care generally did not include screening mammography, or only at specific age thresholds.

The two Canadian trials enrolled volunteers who underwent a pre-examination with CBE before randomization to the intervention or control groups. The Swedish trials randomized women according to communities. The Age trial recruited women from general practice lists, and the
HIP trial recruited women enrolled in a health insurance plan.

The Gothenburg, Stockholm, Malmö, Swedish Two-County, Age, and HIP trials included DCIS in their breast cancer case reporting, while the Canadian trials included only invasive breast cancer in the latest update. All of the trials provided information on the stage, size, or lymph node involvement of cases; however, these outcomes were reported differently across the trials using various descriptions and levels of severity.

Trials differed in their methods of accrual of breast cancer cases and deaths, influencing the analysis of outcomes. Two methods are provided in this report to help explain discrepancies between estimates (Figure 4). The long case accrual method counts all of the breast cancer cases contributing to breast cancer deaths diagnosed during the screening intervention period plus the followup period. This method has been referred to as the “followup” method of analysis by some investigators. While this method includes the most cases, it has the potential to dilute a true benefit because participants from the control group are also screened after the study intervention period ends.

The short case accrual method includes only deaths occurring among cases of breast cancer diagnosed during the screening intervention period, and in some trials, within an additional defined case accrual period. This has been referred to as the “evaluation method” of analysis by some investigators. This method always involves the evaluation of fewer breast cancer cases for mortality outcomes because the duration of case accrual is shorter than for the long case accrual period. This method reduces the risk of contamination in the control group after the screening phase of a trial is completed, but in the absence of concurrent screening, it can introduce bias.

The applicability of the screening trials to current populations and practice has likely decreased over time. All of the trials were conducted in the past when imaging technologies and breast cancer therapies were markedly different than today. Only the HIP trial enrolled women in the United States, however, this trial began 50 years ago. Only women in the Canadian and Age trials, and some women in the Malmö trial, had access to current adjuvant chemotherapies for breast cancer.

In general, women who enroll in trials and attend screening interventions differ from those who do not, underscoring the importance of intention-to-screen analysis to evaluate outcomes. Two trials (HIP, Stockholm) evaluated the differences between women randomized to the intervention group who chose to be screened (attendees) compared with those who did not. In these trials, attendees had higher risks of breast cancer and lower risks of all-cause mortality than non-attendees. In the Canadian trials that recruited volunteers from several communities, participants were more educated, had lower parity, and had overall higher risks of breast cancer compared with the general population. These findings indicate that women at higher risk of breast cancer but lower risk of all-cause mortality may choose to participate in screening. These are important differences that could influence outcomes.
Key Question 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?

Summary

Randomized Trials of Screening

- Updated results from the CNBSS-1, CNBSS-2, Age, and Swedish Two-County Study trials provided breast cancer mortality outcomes with longer followup than the previous review.
- For women age 39 to 49 years, a meta-analysis of trials comparing mammography screening with nonscreening indicated a combined RR of 0.88 (95% confidence interval [CI], 0.73 to 1.003; 9 trials) using the long case accrual method; and 0.84 (95% CI, 0.70 to 1.002; 9 trials) with short case accrual. The absolute mortality reduction (deaths prevented) with screening was 4 per 10,000 women over 10 years.
- For age 50 to 59 years, the combined RR was 0.86 (95% CI, 0.68 to 0.97; 7 trials) with long case accrual; and 0.86 (95% CI, 0.69 to 1.007; 7 trials) with short case accrual. The absolute mortality reduction with screening was 5 to 8 per 10,000 women over 10 years.
- For age 60 to 69 years, the combined RR was 0.67 (95% CI, 0.54 to 0.83; 5 trials) with long case accrual; and 0.67 (95% CI 0.55 to 0.91; 5 trials) with short case accrual. The absolute mortality reduction (deaths prevented) with screening was 12 to 21 per 10,000 women over 10 years.
- Breast cancer mortality for women age 70 to 74 years was not statistically significantly different between randomized groups in the screening trials, but estimates were limited by low numbers of events from trials that had smaller sample sizes of women in this age group.
- All-cause mortality did not differ between randomized groups in meta-analyses of trials, regardless of whether trials were analyzed in combined or separate age groups.
- No RCTs evaluated breast cancer mortality or all-cause mortality outcomes on the basis of risk factors besides age.
- There are no head-to-head trials of different screening intervals and existing trials do not provide enough information to determine the specific effects of screening intervals.

Observational Studies

- Observational studies of the effectiveness of population-based mammography screening on breast cancer mortality reported a wide range of reductions in breast cancer death. Most studies were conducted in Europe or the United Kingdom and included women age 50 to 69 years.
- Meta-analyses from recent reviews from the EUROSCREEN Working Group indicated 25 to 31 percent mortality reduction for women invited to screening in the screening programs. This compares to 19 to 22 percent reduction for women age 50 to 69 years in the meta-analysis of screening RCTs that used intention-to-screen analysis.
• The only U.S observational study of breast cancer mortality reduction is a record review that indicated no differences in breast cancer deaths between screened versus non-screened women older than age 80 years.91
• A large fair-quality study of the Mammography Screening of Young Women Cohort in Sweden indicated reduced risk for breast cancer deaths for women age 40 to 49 years invited to screening compared with women not invited (RR 0.74; 95% CI, 0.66 to 0.83).
• An observational study of Canadian women age 40 to 79 comparing screening program participants versus nonparticipants indicated reduced breast cancer mortality of 40 percent among participants.
• Two observational studies of screening intervals indicated no breast cancer mortality differences between annual and biennial screening for women 50 years or older, or between annual and triennial screening among women age 40 to 49 years.

Evidence

Previous Reports

The 2002 evidence review for the USPSTF included a meta-analysis of the eight published RCTs of mammography screening and breast cancer mortality that were rated fair-quality.92,93 For all age groups combined, results of the meta-analysis indicated a RR for breast cancer mortality of 0.84 (95% credible interval [CrI], 0.77 to 0.91) for women randomly assigned to screening over 14 years of followup. For women age 40 to 49 years specifically, results indicated a RR of 0.85 (95% CrI, 0.73 to 0.99), while for women age 50 years and older, results indicated a RR of 0.78 (95% CrI, 0.70 to 0.87).

The 2009 evidence review for the USPSTF included new results from the Age trial and updated results from the Gothenburg trial in addition to the previous trials, and provided meta-analysis estimates for breast cancer mortality according to four age groups.44,94 For women age 39 to 49 years, the combined RR was 0.85 (95% CrI, 0.75 to 0.96); for 50 to 59 years, 0.86 (95% CrI, 0.75 to 0.99); for 60 to 69 years, 0.68 (95% CrI, 0.54 to 0.87); and for 70 to 74 years, 1.12 (95% CrI, 0.73 to 1.72).

Previous evidence reviews for the USPSTF did not address the effectiveness of screening in reducing all-cause mortality, or how mortality reduction differs by risk factors and screening intervals.

New Studies

Breast Cancer Mortality

RCTs with long case accrual methods. Seven RCTs provided breast cancer mortality outcomes by age using long case accrual methods. These included the Swedish Two-County (Kopparberg and Östergötland),83 Age,79 Gothenburg,82 HIP,95 and Canadian (CNBSS-1 and CNBSS-2)69 trials. The Malmö I, Malmö II, and Stockholm trials reported breast cancer mortality outcomes by age using only short case accrual.87 However, these results were included in the combined meta-analysis because they are the most inclusive results available. Quality ratings of these trials
are described in previous reports.\textsuperscript{92,93,94} Across all trials with long case accrual, 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 followup time from 11.2 to 21.9 years.

For women age 39 to 49 years, a meta-analysis of nine RCTs using the longest case accrual available indicated a combined RR of 0.88 (95% CI, 0.73 to 1.00; $I^2=25.4\%$; $p=0.218$; Figure 5).\textsuperscript{69,82,83,86,87,95} The CIs of all nine trials crossed 1.0 as did the combined estimate.

For age 50 to 59 years, a meta-analysis of seven trials using the longest case accrual available indicated a combined RR of 0.86 (95% CI, 0.68 to 0.97; $I^2=38.0\%$; $p=0.139$), consistent with a statistically significantly lower death rate in the screening group.\textsuperscript{69,82,83,87,95} Estimates from the Kopparberg\textsuperscript{83} and Stockholm\textsuperscript{87} trials indicated statistically significant differences between randomized groups favoring screening, while the CIs from the five other trials crossed 1.0.

For age 60 to 69 years, a meta-analysis of five trials using the longest case accrual available indicated a combined RR of 0.67 (95% CI, 0.54 to 0.83; $I^2=0\%$; $p=0.739$), consistent with a statistically significantly lower death rate in the screening group.\textsuperscript{83,87,95} In this age group, estimates from three Swedish trials (Kopparberg,\textsuperscript{83} Östergötland,\textsuperscript{83} and Malmö I\textsuperscript{87}) indicated statistically significant differences between randomized groups favoring screening, while the CIs from the two other trials crossed 1.0. Combining results across the two age groups of women age 50 to 69 years indicated a RR of 0.78 (95% CI, 0.68 to 0.90; $I^2=41.0\%$; $p=0.118$).

Only three Swedish trials, Östergötland,\textsuperscript{83} Kopparberg,\textsuperscript{83} and Malmö I,\textsuperscript{87} provided outcomes for women age 70 to 74 years. The numbers of events in these trials were much lower than for other age groups, and none of the trials indicated statistically significant differences between randomized groups. A meta-analysis of the three trials using the longest case accrual available indicated a combined RR of 0.80 (95% CI, 0.51 to 1.28; $I^2=0\%$; $p=0.962$).

A sensitivity analysis that included results of a combined analysis of the Swedish trials (Malmö I, Malmö II, Stockholm, Östergötland, Gothenburg, Stockholm) that used a long case accrual (“followup”) method\textsuperscript{87} indicated reduced point estimates that diminished the effect of screening, although the statistical significance of the estimates did not change.

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 5). Using RR$s$ from the long case accrual meta-analysis, the numbers of deaths reduced (prevented) included 4.1 (95% CI, -0.1 to 9.3) for age 39 to 49 years; 7.7 (95% CI, 1.6 to 17.2) for age 50 to 59 years; 21.3 (95% CI, 10.7 to 31.7) for age 60 to 69 years; and 12.5 (95% CI, -17.2 to 32.1) for age 70 to 74 years. Absolute reduction for the combined group of women age 50 to 69 was 12.5 (95% CI 5.9 to 19.5).

**RCTs with short case accrual methods.** Meta-analysis estimates from trials with short case accrual methods differed only slightly from those with long case accrual (Figure 6). 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 followup time from 10.7 to 25.7 years. Including the same trials as the previous analysis, but with short case accrual, the combined RR for women age 39 to 49 years was 0.84 (95% CI, 0.70 to 1.002; $I^2=35.8\%$; $p=0.143$; 9
The Gothenburg trial was the only trial with statistically significant differences between groups.\textsuperscript{82} Results for age 50 to 59 years indicated a RR of 0.86 (95% CI, 0.69 to 1.00; \(I^2=33.9\%\); \(p=0.182\); 7 trials), and only the Stockholm trial reported statistically significant differences between groups.\textsuperscript{83,87} Results for age 60 to 69 and 70 to 74 years differed slightly from the previous analysis (60 to 69 years; RR 0.67; 95% CI 0.55 to 0.91; \(I^2=0\%\); \(p=0.476\); 5 trials; and 70 to 74 years; RR 0.90; 95% CI, 0.46 to 1.78; \(I^2=0\%\); \(p=0.923\); 3 trials). Combining results across the two age groups of women age 50 to 69 years, indicated a RR of 0.81 (95% CI, 0.69 to 0.95; \(I^2=43.7\%\); \(p=0.114\)).

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 5). Using RRs from the short case accrual meta-analysis, the numbers of deaths reduced (prevented) included 3.5 (95% CI, -0.1 to 7.4) for age 39 to 49 years, 4.5 (95% CI, -0.2 to 9.8) for age 50 to 59 years, 12.1 (95% CI, 3.4 to 20.7) for age 60 to 69 years, and 12.2 (95% CI, -37.7 to 26.9) for age 70 to 74 years. Absolute reduction for the combined group of women age 50 to 69 years was 6.1 (95% CI 1.2 to 10.9).

**Observational studies.** Observational studies of mammography screening provide additional information about screening effectiveness in contemporary populations and settings. However, observational studies are subject to important biases that limit their use in determining effectiveness. Most importantly, they lack comparability of comparison groups that is only attainable through randomization.

Recent comprehensive systematic reviews of observational studies summarize most of the relevant research.\textsuperscript{96-99} Included studies were designed as time-trend, incidence-based mortality, or case-control studies. Time-trend studies compare changes in breast cancer mortality among populations in relation to the introduction of screening. Incidence-based mortality studies compare mortality rates of women screened or invited to screen with women not screened or invited. To reflect the incidence of breast cancer, rather than prevalence, these studies include only breast cancer cases diagnosed during a specific time period that follows the initial screen. Case-control studies compare histories of screening between women dying of breast cancer with women not dying of breast cancer. Examples of limitations of these specific study designs include incorrect assumptions for comparison groups in time-trend studies, high risk of lead and length time bias in incidence-based mortality studies, and self-selection bias in case-control studies. Additional limitations are described in Table 6.

Three good-quality reviews were recently conducted by the EUROSCREEN Working Group to assess the effectiveness of population-based mammography screening on breast cancer mortality (Appendix B1).\textsuperscript{96-98} Inclusion criteria included studies with original data from population-based screening programs in Europe and the United Kingdom that reported breast cancer mortality outcomes; were published in English; included women age 50 to 69 years; evaluated current screening programs; and were designed as time-trend, incidence-based mortality, or case-control studies. Studies with overlapping data or data that were updated by newer results were not included. Although quality criteria were not prespecified, the studies appeared to undergo critical review according to design-specific factors. However, individual studies were not given quality.
ratings. Studies included in these reviews are listed in Appendix A8.

A EUROSCREEN review evaluated 12 time-trend studies reporting changes in breast cancer mortality in relation to the introduction of screening. These studies described trends over time or evaluated change using regression analysis. No combined estimates of effectiveness were provided because of dissimilarities of comparisons and outcome measures. Five studies reporting outcomes as reductions per year indicated breast cancer mortality reductions of 1 to 9 percent per year for approximately 10 years after the introduction of screening (i.e., 10% to 90%).

Seven studies reporting before/after changes indicated 0 to 36 percent reductions in mortality after screening was introduced compared with before screening. Three of these studies that were considered to have adequate followup reported mortality reductions ranging from 28 to 35 percent.

Another EUROSCREEN review included 20 incidence-based mortality studies that evaluated breast cancer mortality rates in relationship to screening. The least biased studies estimated breast cancer mortality from a cohort of women not invited for screening, or from historical and current control groups; and used long case accrual periods that were the same as the study followup periods. A meta-analysis of these studies indicated a RR of 0.75 (95% CI, 0.69 to 0.81; p=0.23; 7 studies) for invitation to screening; and 0.62 (95% CI, 0.56 to 0.69; p=0.40; 7 studies) for actual screening.

The third EUROSCREEN review included eight case-control studies that provided odds ratios (ORs) for breast cancer mortality adjusted for self-selection bias using various methods. A meta-analysis of studies indicated an OR of 0.69 (95% CI, 0.57 to 0.83; p=0.005; 7 studies) for invitation to screening; and 0.52 (95% CI, 0.42 to 0.65; p=0.17; 7 studies) for actual screening.

A good-quality systematic review conducted outside of the EUROSCREEN Working Group included time-trend, cohort, and hybrid studies. Hybrid studies were defined as studies that identified a cohort, but used population-based data on mammography exposure. Studies were restricted to those with women age 50 to 69 years that captured over 10 years of screening experience. Several studies included in this review were also included in the EUROSCREEN reviews. Study quality was evaluated by prespecified criteria that included concepts of the USPSTF criteria and emphasized control groups, adjustment for potential confounders, and ascertainment of mortality outcomes. Of 17 studies meeting inclusion criteria and rated fair-quality, five reported RR reductions for breast cancer death of 0 to 12 percent; eight reported 13 to 33 percent; and four reported more than 33 percent, although not all results reached statistical significance.

The results of these systematic reviews indicated a wide range of estimates of breast cancer mortality reduction with screening for women age 50 to 69 years. Meta-analyses from the EUROSCREEN reviews indicated 25 to 31 percent mortality reduction for women invited to screening in the screening programs. In comparison, the meta-analysis of screening RCTs using intention-to screen analysis for women age 50 to 69 years indicated reductions of 19 to 22 percent, as described in the previous section of this report.
Six additional studies were not included in the published systematic reviews described above because they were published in 2011 or later, included women in countries outside Europe and the United Kingdom, or focused on ages older or younger than 50 to 69 years (Table 7). The only U.S study was a record review of older women who died of breast cancer. Results indicated no differences in breast cancer deaths between screened versus non-screened women older than age 80 years.91

One study included only women in their 40s. A large fair-quality study of the Mammography Screening of Young Women Cohort in Sweden indicated reduced risk for breast cancer deaths for women age 40 to 49 years invited to screening compared with women not invited (RR 0.74; 95% CI, 0.66 to 0.83).123 The estimated NNS during a 10-year period (corresponding to about 6 mammography episodes) to save 1 life was calculated as 1252 women (95% CI, 958 to 1915 women).123

A study of over 2 million women age 40 to 79 in Canada compared screening program participants versus nonparticipants.124 Results were expressed as standardized mortality ratios (i.e., the ratio of the observed breast cancer mortality of screening participants to province-specific breast cancer mortality based in nonparticipant incidence and survival rates).124 Results indicated reduced breast cancer mortality of 35 to 44 percent that varied by age. Although the analysis considered the influence of self-selection bias using historical trend data for women age 35 to 39, the validity of this approach is unclear.

Additional studies provided updated data from screening programs in Norway125,126 and the Netherlands127 with results consistent with the EUROSCREEN report showing reduced mortality with screening for women age 50 to 69 years.

**All-Cause Mortality**

All included RCTs of mammography screening reported all-cause mortality outcomes. However, not all trials reported them according to age groups, and the two Canadian trials reported results by combining age groups (40 to 49 years and 50 to 59 years) as one trial. Results reflecting the longest followup times available for each trial were selected for inclusion in the meta-analysis.

For combined age groups, a meta-analysis of nine RCTs indicated a combined RR of 0.99 (95% CI, 0.97 to 1.003; $I^2=0$%; $p=0.577$, Figure 7).69,79,87,128 Results were similar for each age group (Figure 8), including age 39 to 49 years (RR 0.99; 95% CI, 0.94 to 1.06; $I^2=0$%; $p=0.478$; 7 trials); 50 to 59 years (RR 1.02; 95% CI, 0.94 to 1.10; $I^2=0$%; $p=0.588$; 3 trials); 60 to 69 years (RR 0.97; 95% CI, 0.90 to 1.04; $I^2=72.4$%; $p=0.057$; 2 trials); and 70 to 74 years (RR 0.98; 95% CI, 0.86 to 1.14; $I^2=72.4$%; $p=0.057$; 2 trials).

**Breast Cancer–Specific and All-Cause Mortality Differences by Risk Factors and Screening Intervals**

Screening trials did not provide results according to risk factors other than age. No head-to-head comparisons of trials by screening intervals are available. The HIP, Age, and Canadian trials used mammography screening intervals of 12 months, and none showed age-specific mortality
reductions. The Swedish Two-County trial had screening intervals ranging from 24 to 36 months that varied by age group, and reported breast cancer mortality reductions for age 50 to 69 years. However, these trials differed by many other factors (inclusion, randomization, adherence, etc.) and they did not provide enough information to determine the specific effects of screening intervals.

Observational studies provide additional information about screening intervals (Table 7). A time-trend study of 658,151 Canadian women age 40 to 79 years compared breast cancer mortality rates before and after the change from annual to biennial screening for women 50 years or older, while annual screening remain unchanged for age 40 to 49 years. Results indicated no significant reductions for age 40 to 49 or 50 years and older. A registry-based study in Finland indicated no breast cancer mortality differences between annual and triennial screening among women age 40 to 49 years.

Key Question 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?

Summary

RCTs

- The RCTs of mammography screening provided several measures of intermediate breast cancer outcomes. However, most comparisons between screening and control groups using these categories provided differences between the two groups in relatively early stages of disease, rather than advanced stages.
- Combining estimates based on definitions corresponding to Stage II disease or higher (Stage II+, size ≥20 mm, 1+ positive lymph node) in a meta-analysis indicated no significant reductions in advanced disease for women age 39 to 49 or 50 years and older.
- When thresholds were defined by the most severe disease categories available from the trials (Stage III + IV disease, size ≥50 mm, 4+ positive lymph nodes), meta-analysis indicated no reductions for age 39 to 49 years (RR 0.98; 95% CI, 0.74 to 1.37); but reduced risk of advanced cancer in the screening group for age 50 years and older (RR, 0.62; 95% CI, 0.46 to 0.83).
- In a Cochrane review that included five screening RCTs, women randomized to screening were significantly more likely to have surgical therapy (mastectomies, lumpectomies) and radiation therapy, and less likely to have hormone therapy than controls. Use of chemotherapy was similar between groups.
- No RCTs evaluated the incidence of advanced breast cancer outcomes and treatment on the basis of risk factors or screening intervals.
Observational Studies

- Six observational studies compared advanced breast cancer outcomes between women in populations participating in screening versus nonparticipating. Of these, two studies indicated statistically significantly more Stage III and IV breast cancer among unscreened women; three reported more lymph node positive disease; and three reported more tumors greater than 20 mm in size.
- Four case series studies indicated less extensive survey, such as fewer total mastectomies and more breast conservation therapies, and less chemotherapy among women who had previously had screening mammography compared with those who did not, but these studies included women with DCIS and early stage cancer as well as advanced cancer.
- An analysis of BCSC data indicated a lower proportion of Stage III + IV disease among women age 40 to 49 years screened annually versus biennially, but not for women age 50 to 59 years.
- A second analysis of BCSC data indicated that women age 40 to 49 years with extremely dense breasts had increased risks for advanced stage cancer (IIB+) and large-size tumors (>20 mm) with biennial compared with annual screening. Differences were not significantly different for positive lymph nodes, other density categories, other age groups, or between biennial and triennial screening.

Evidence

Previous Reports

Previous evidence reviews for the USPSTF did not address this question.

New Studies

Incidence of Advanced Breast Cancer

RCTs. Intermediate outcomes of screening trials can be evaluated to determine if screening reduces the risk for advanced breast cancer, thereby leading to better prognosis and potentially less aggressive treatment and morbidity. The RCTs of mammography screening provided several measures of intermediate outcomes for screening and control groups. The most commonly used measures included clinical stage (Stage 0 to IV), number of involved lymph nodes (0, 1 to 3, 4+), and tumor size (mm), although these measures varied across trials. Most comparisons between screening and control groups using these categories provided differences between the two groups in relatively early stages of disease, rather than advanced stages.

A published analysis of trials defined advanced breast cancer as Stage II disease or higher, size 20 mm or greater, or having one or more positive lymph nodes (Table 8). These outcomes are all consistent with Stage IIA disease (i.e., localized) or higher according to the AJCC TNM system. Combining estimates based on these definitions of advanced cancer in a meta-analysis produced a RR for women age 39 to 49 years of 0.90 (95% CI, 0.79 to 1.04; I²=23.1%; p=0.267; 5 trials), and for age 50 years and older, 0.85 (95% CI, 0.65 to 1.13; I²=80.5%; p=0.002;
indicating no statistically significant overall differences between the screening and control groups.

To evaluate these relationships using a higher level of disease to define advanced breast cancer, thresholds were redefined to the most severe disease categories available from the trials, recognizing that these definitions do not represent equivalent disease stages. These include Stage III + IV disease (i.e., regional + metastatic), size 50 mm or greater, or having four or more positive lymph nodes. Combining estimates based on these definitions of advanced cancer in a meta-analysis indicated no difference for women age 39 to 49 years (RR 0.98; 95% CI, 0.74 to 1.37; $I^2=0\%$; p=0.556; 4 trials); but reduced risk of advanced cancer in the screening group for age 50 years and older (RR, 0.62; 95% CI, 0.46 to 0.83; $I^2=0\%$; p=0.692; 3 trials; Figure 10).

Observational studies. Although many observational studies have been published comparing characteristics of breast cancer diagnoses between screened and unscreened women, most provide comparisons between screen-detected and nonscreen-detected cases. This approach categorizes all cancer cases identified outside of a screening mammogram as nonscreen-detected, even though a woman may have had prior screening mammography. This type of comparison does not provide accurate estimates of the effectiveness of participation in a screening program compared with nonparticipation. Instead, comparisons between rates of advanced breast cancer outcomes between women in populations participating in screening versus nonparticipating would more appropriately address this Key Question.

Six case series studies compared advanced breast cancer outcomes for women who had previous mammography screening with those who did not (Table 9). These studies were based on screening populations from the Malmö trial, Kaiser Permanente, and screening programs in the United Kingdom, Denmark and Sweden, Spain, and Canada.

Two studies indicated statistically significantly more Stage III and IV breast cancer among unscreened women, three reported more lymph node positive disease, and three reported more tumors greater than 20 mm in size. A study of 242 women age 42 to 49 years at Kaiser Permanente found no statistically significant differences in stage between screened and nonscreened women.

Treatment-Related Morbidity

Although outcomes related to treatment are reported by some of the screening trials, their interpretation and application to current practice is problematic. Treatment approaches have changed over time, are subject to local practice standards, and increasingly involve patient choices.

A Cochrane review compared treatments between randomized groups in five screening trials providing these outcomes, including the CNBSS-1, CNBSS-2, Malmö, Kopparberg, and Stockholm trials. In this analysis, women randomized to screening were significantly more likely to have surgical therapy, analyzed as mastectomies and lumpectomies combined (RR 1.35; 95% CI, 1.26 to 1.44; $I^2=0\%$; p=0.80; 5 trials) and mastectomies alone (RR 1.20; 95% CI, 1.11
to 1.30; $I^2=0\%; p=0.86; 5$ trials). These women were also more likely to have radiation therapy (RR 1.32; 95% CI, 1.16 to 1.50; $I^2=0\%; p=0.36; 2$ trials), and less likely to have hormone therapy (RR 0.73; 95% CI, 0.55 to 0.96; $I^2=78\%; p=0.03; 2$ trials). Use of chemotherapy was similar between groups (RR 0.96; 95% CI, 0.78 to 1.19; $I^2=71\%; p=0.06; 2$ trials).

Four case series studies compared breast cancer treatments for women who had previous mammography screening with those who did not, but these studies included women with DCIS and early stage cancer as well as advanced cancer (Table 10). Studies also provided information on advanced cancer outcomes described above, and were based on screening populations from Kaiser Permanente, and screening programs in Denmark and Sweden, Spain, and Canada. Results indicated statistically significantly less extensive surgery, such as fewer total mastectomies and more breast conservation therapies, and less chemotherapy among women who had previously had screening mammography.

**Differences by Risk Factors and Screening Intervals**

Five observational studies compared advanced breast cancer outcomes by screening intervals (Table 9), including four studies based on BCSC data. A recent analysis of data from 4,492 women in the BCSC compared annual with biennial mammography screening and the adjusted proportion of cancer stage at diagnosis. Results indicated a lower proportion of Stage III + IV disease among women age 40 to 49 years screened annually versus biennially (10.1% vs. 14.0%; adjusted difference 4.8%; 95% CI, 1.3% to 8.4%), but not among women age 50 to 59 years. An older study of 7,840 women in the BCSC indicated no differences between annual and biennial screening for detecting Stage III + IV cancer or tumor size greater than 20 mm among women age 40 to 89 years.

A separate analysis of BCSC data compared annual with biennial and triennial mammography screening and risks for advanced stage disease (Stage IIB+), large tumor size (>20 mm), and positive lymph nodes. Results indicated that women age 40 to 49 years with extremely dense breasts had increased risks for advanced stage cancer (OR 2.39; 95% CI, 1.06 to 3.39) and large tumors (OR 2.39; 95% CI, 1.37 to 4.18) with biennial compared with annual screening. Differences were not statistically significantly different for positive lymph nodes, other density categories, other age groups, or between biennial and triennial screening. Another BCSC study reported no statistically significant differences in stage, tumor size, or lymph node involvement for average weight, overweight, and obese women screened annually compared with biennially.

A study based on data from the Vermont Breast Cancer Surveillance System also reported no differences in cancer stage, size, or lymph node status between women screened annually compared with biennially.
Key Question 3. How Does the Effectiveness of Routine Breast Cancer Screening in Reducing Breast Cancer–Specific and All-Cause Mortality Vary by Different Screening Modality?

Summary

- RCTs of mammography with or without CBE do not compare relative mortality reduction across different modalities.
- No study of tomosynthesis, ultrasound, or MRI address this question.

Key Question 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?

Summary

- Cancer detection rates were higher, but there were no differences in tumor size, stage, or node status between women screened with tomosynthesis and digital mammography and with those receiving mammography alone in two case series studies.
- No other studies evaluated the effectiveness of CBE, ultrasound, or MRI in reducing the incidence of advanced breast cancer or treatment related morbidity.

Evidence

Previous Reports

Previous evidence reviews for the USPSTF did not address this question.

New Studies

Two case series studies comparing digital mammography versus tomosynthesis and digital mammography reported detection rates by cancer stage using various categories of cancer staging (Table 11). A study of patients seen at a multisite community-based breast center in the United States evaluated diagnostic outcomes of 18,202 women receiving mammography and 10,878 receiving mammography and tomosynthesis. Results indicated no differences in cancer size, stage, or node status. A second case series of 12,631 women age 50 to 69 years in Norway also found no differences between groups for tumor size or node status but found a 27 percent adjusted increase in cancer detection rates (p=0.001) with the addition of tomosynthesis.
Key Question 5. What Are the Harms of Routine Mammography Screening, and How Do They Differ by Age, Risk Factor, and Screening Interval?

Summary

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

- Data from the BCSC for regularly screened women using digital mammography based on results from a single screening round indicated:
  - False-positive mammography rates were highest among women age 40 to 49 years (121.2 per 1,000 women; 95% CI 105.6 to 138.7) and declined with age; rates of false-negative results tended to increase with age, but were not statistically significantly different across age groups.
  - Rates of recommendations for additional imaging were highest among women age 40 to 49 years (124.9 per 1,000 women; 95% CI 109.3 to 142.3) and decreased with age, while rates of recommendations for biopsy did not differ between age groups.
  - For every case of invasive breast cancer detected by mammography screening in women age 40 to 49 years, 464 women had screening mammography, 58 were recommended for additional imaging, and 10 were recommended for biopsies. These estimates declined with age.
  - Results did not differ by time since last mammography screening regardless of whether broad or narrow estimates of one versus two years were used.
  - Family history of breast cancer, high breast density, and previous benign breast biopsy were associated with higher rates of false-positive and false-negative results and recommendations for additional imaging and biopsy across most age groups. Premenopausal status, use of menopausal hormone therapy, and lower BMI were associated with some of the outcomes for specific age groups only.
  - Rates for all outcomes were lowest for women with almost entirely fat breasts, and highest for women with heterogeneously dense breasts or for those in the combined category of heterogeneous and extreme density.

- Published data from the BCSC using film and digital mammography provided 10-year cumulative rates.
  - Rates of false-positive mammography results were 61 percent for annual and 41 percent for biennial screening, while rates of false-positive biopsy were 7 to 9 percent for annual and 5 to 6 percent for biennial screening. Women older than age 50 years had higher false-positive biopsy rates.
  - Rates of false-positive mammography results and biopsy were highest among women receiving annual mammography, those with heterogeneously dense or extremely dense breasts, and those either 40 to 49 years old or who used combination hormone therapy.

Overdiagnosis

- A meta-analysis of three RCTs, a systematic review of 13 observational studies, and 18
individual studies of overdiagnosis were identified for the current update. Studies of overdiagnosis were primarily based on screening trials, screening programs and registries, or modeled data. Studies differed by their characteristics, methods, and measures. These differences influenced their estimates of overdiagnosis, limited comparisons, and prohibited combined estimates.

- The Malmö I and Canadian National Breast Screening Study (CNBSS-1 and CNBSS-2) trials provide data with reduced bias for estimates of overdiagnosis because they did not provide screening of controls at the end of the trial, had randomized comparison groups, and followup times extended sufficiently beyond the screening period to differentiate earlier diagnosis from overdiagnosis. Combined results indicated 10.7 percent (long case accrual method) to 19.0 percent (short case accrual method) overdiagnosis for invasive cancer + DCIS.
- Data from RCTs where women in the control groups were offered screening at the end of the screening periods are susceptible to over- or underestimating overdiagnosis. Two new publications from these RCTs indicate no or minimal overdiagnosis.
- Unadjusted estimates from 13 observational studies included in the EUROSCREEN review indicated overdiagnosis rates ranging from 0 to 54 percent. For six studies that adjusted overdiagnosis estimates for breast cancer risk and lead time, rates varied from 1 to 10 percent.
- Additional observational studies not included in the EUROSCREEN review reported overdiagnosis estimates of 3 to 50 percent, with most between 14 to 25 percent.
- Although several statistical models of overdiagnosis have been published, these studies have been less acceptable to guideline development groups because of the many assumptions that were used to construct them. Models indicated estimates ranging from 0.4 to 50 percent.

Anxiety, Distress, and Other Psychological Responses

- Women with false-positive results were more distressed than women with normal screening results, particularly those who had biopsies, FNA, and early recall.
- Women with false-positive results had more anxiety, psychological distress, and breast cancer specific worry after screening compared with those with normal screening results in most studies. Anxiety improved over time for most women, but persisted for over 2 years for some.
- Two studies reported that women with false-positive results were less likely to return for their next mammogram; two other studies reported no differences; however, when women were given letters tailored to their last screening result they were more likely to re-attend.
- Results of studies of anxiety and depression are mixed. Some studies indicate that women with false-positive results have more anxiety and depression than those with normal screening results, particularly among non-white women, but other studies show no differences.

Radiation Exposure

- Models calculate the number of deaths due to radiation induced cancer using estimates for digital mammography is between 2 per 100,000 in women age 50 to 59 years screened biennially, and up to 11 per 100,000 in women ages 40 to 59 years screened annually.
Pain During Procedures

- Although many women may experience pain during mammography (1% to 77%), the proportion of those experiencing pain who do not attend future screening varies (11% to 46%).

Evidence

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

Previous Reports

Data from the BCSC that was based on a single screening round and included film and digital mammography indicated that false-positive mammography results were common in all age groups. The rate was highest among women age 40 to 49 years (97.8 per 1,000 women per screening round) and declined with each subsequent age decade.

The rate of false-negative mammography results was lowest among women age 40 to 49 years (1.0 per 1,000 women per screening round) and increased slightly with subsequent age decades. Rates of additional imaging were highest among women age 40 to 49 years (84.3 per 1,000 women per screening round) and decreased with age. Biopsy rates were lowest among women age 40 to 49 years (9.3 per 1,000 women per screening round) and increased with age. The BCSC data indicated that for every case of invasive breast cancer detected by mammography screening in women age 40 to 49 years, 556 women had screening mammography, 46 to 48 additional diagnostic imaging, and five to eight biopsies. These numbers declined with age for mammography and additional imaging, and only slightly for biopsies.

The cumulative risk for false-positive mammography results was reported in published studies as 21 to 49 percent after 10 mammography examinations for women in general,\textsuperscript{148-150} and up to 56 percent for women age 40 to 49.\textsuperscript{149} For all ages, the cumulative risk of a false-positive biopsy after 10 screening mammograms was calculated as 19 percent of all women screened.\textsuperscript{149}

New Studies

BCSC Data.

\textit{Differences by age.} Data for regularly screened women based on results from a single screening round using digital mammography indicated that false-positive screening mammography results were common in all age groups (Table 12). The rate was highest among women age 40 to 49 years (121.2 per 1,000 women; 95% CI 105.6 to 138.7) and declined with age (p<0.001). Rates of false-negative mammography results tended to increase with age, but were not statistically significantly different across age groups, and ranged from 1.0 to 1.5 per 1,000 women.

In current practice, women with an initial positive mammography result are recommended for
additional diagnostic imaging as a second step in the screening process. In the BCSC data, rates of recommendations for additional imaging were highest among women age 40 to 49 years (124.9 per 1,000 women; 95% CI 109.3 to 142.3) and decreased with age (p<0.001). Rates of recommendations for biopsy were not statistically significantly different across age groups, and ranged from 15.6 to 17.5 per 1,000 women.

Rates of invasive breast cancer were lowest among women aged 40 to 49 years (2.2 per 1,000 women; 95% CI 1.8 to 2.6) and increased across age groups (p<0.001). Rates of DCIS were also lowest among women aged 40 to 49 years (1.6 per 1,000 women; 95% CI 1.3 to 1.9) and increased with age (p=0.05). Women aged 70 to 79 had the highest rates of invasive cancer (7.2 per 1,000 women; 95% CI 6.4 to 8.1) and DCIS (2.3 per 1,000 women; 95% CI 1.7 to 3.0), and the yield of screening was more favorable for older women. For every case of invasive breast cancer detected by mammography screening in women aged 40 to 49 years, 464 women had mammography, 58 were recommended for additional imaging, and 10 were recommended for biopsies. In comparison, for women aged 70 to 79, for every case of invasive breast cancer detected by screening, 139 women had mammography, 11 were recommended for additional imaging, and 3 were recommended for biopsies.

**Differences by time since last mammography screening.** Rates of false-positives, false-negatives, and recommendations for additional imaging did not differ in comparisons of times since last mammography screening regardless of interval durations (9 to 18 versus 19 to 30 months; 11 to 14 versus 23 to 26 months) (Table 13). Biopsies were recommended at a higher rate for women aged 60 to 69 years who had their last mammogram 23 to 26 months previously compared to 11 to 14 months (18.8 versus 15.2 per 1,000 women; p=0.03).

Additional published BCSC data about screening intervals indicated that sensitivity, recall rates, and cancer detection rates increased as the months since previous mammography increased, whereas specificity decreased.151

**Differences by risk factors.** Rates of false-positive mammography results were statistically significantly higher for women with specific risk factors compared with women without them (Table 14). These included women with first-degree relatives with breast cancer compared with no relatives for women aged 40 to 69 years. Women with heterogeneously dense breasts had higher false-positive rates than those with almost entirely fat and scattered fibroglandular densities, or extremely dense breasts, for all ages except aged 80 to 89 years. Rates were also higher among women with previous benign breast biopsies for aged 40 to 79 years. Comparisons based on race and ethnicity indicated the lowest rates among Asians for all age groups.

Premenopausal women had the highest false-positive rates for women aged 40 to 59 years compared with perimenopausal and postmenopausal women. Women using menopausal hormone therapy had the highest rates for aged 70 to 79 years, while comparisons for other age groups were not statistically significant. Women with lower body mass (BMI <30) had higher false positive rates for aged 40 to 59 years.

Rates of false-negative results were higher for women with first-degree relatives with breast cancer for aged 40 to 79 years, although results were of borderline statistical significance for
aged 50 to 69 years (Table 15). Women with almost entirely fat and scattered fibroglandular densities had lower rates than those with other types of breast density for ages 40 to 69 years. Rates were higher among women with previous benign breast biopsies for aged 50 to 89 years, and women with lower body mass (BMI <30) for ages 50 to 59 years. Other comparisons between groups were not statistically significant.

Risk factors associated with differences in rates of recommendations for additional imaging were similar to those for false positive mammography results (Table 16). Rates were highest among women with first-degree relatives with breast cancer for all ages, heterogeneously dense breasts (ages 40 to 79), previous benign breast biopsies (ages 40 to 79), premenopausal status (ages 40 to 50), use of menopausal hormone therapy (age 70 to 79), and lower BMI (ages 40 to 49). Comparisons based on race and ethnicity indicated the lowest rates among Asians for all age groups.

Rates of recommendations for biopsy were statistically significantly higher for women aged 40 to 69 years with first-degree relatives with breast cancer, and for women aged 40 to 79 years with previous breast biopsies (Table 17). Women aged 40 to 59 years with heterogeneously dense or extremely dense breasts had higher rates than women with less dense breasts, while for women aged 60 to 79 years, rates were highest for women with heterogeneously dense breasts only. Higher rates were also associated with premenopausal status for age 50 to 59 years; no current use of oral contraceptives for age 40 to 49 years; lower BMI for age 40 to 49, but higher BMI for age 70 to 79. Other comparisons between groups were not statistically significant.

Rates of false-positives, false-negatives, recommendations for additional imaging, and recommendations for biopsies were lowest for women with almost entirely fat breasts for all ages. False-negative rates were highest for women with extremely dense breasts for all ages, except those aged 60 to 69 years (Table 18). Rates of false-positives, recommendations for additional imaging, and recommendations for biopsies were highest for women with heterogeneously dense breasts or for the combined category of heterogeneously and extremely dense breasts, except for women aged 40 to 49 years where rates of recommendations for biopsies were highest among women with extremely dense breasts.

Two studies published since the 2009 review estimated the cumulative probability of false-positive results after 10 years of mammography screening based on data from the BCSC.143,144 Data collected from film and digital screening mammography performed between 1994 and 2006 indicated that when screening began at age 40 years, the cumulative probability of receiving at least one false-positive mammography result after 10 years was 61 percent (95% CI, 59% to 63%) with annual, and 41 percent (95% CI, 41% to 43%) with biennial screening.143 Estimates were similar when screening began at age 50 years. The cumulative probability of receiving a false-positive biopsy recommendation after 10 years of screening was 7 percent (95% CI, 6% to 8%) with annual versus 5 percent (95% CI, 4% to 5%) with biennial screening for women who initiated screening at age 40 years; and 9 percent (95% CI, 7% to 12%) with annual versus 6 percent (95% CI, 6% to 7%) with biennial for women who began at age 50 years.

A study of BCSC data collected between 1994 and 2008 also evaluated 10-year cumulative
probability estimates for false-positive mammography and biopsy results, but stratified results by age, breast density, and use of menopausal hormone therapy. Rates of false-positive mammography results were highest among women receiving annual mammography that had extremely dense breasts and were either 40 to 49 years old (65.5%) or used combination hormone therapy (65.8%). Rates were lower among women 50 to 74 years receiving biennial or triennial mammography that had scattered fibroglandular densities (39.7% and 21.9%, respectively) or almost entirely fat breasts (17.4% and 12.1%, respectively). These rates were similar regardless of menopausal estrogen use. The highest rates were among women age 40 to 49 years undergoing annual screening that had heterogeneously dense (68.9%) or extremely dense (65.5%) breasts. The highest rates of false-positive biopsy were related to similar characteristics and ranged from 12 to 14 percent.

Overdiagnosis

Previous Reports

A review of eight RCTs of mammography screening and eight additional studies in the previous report provided estimates of overdiagnosis that ranged from non-existent to nearly 50 percent of diagnosed breast cancer cases. Methods for estimating overdiagnosis varied in many ways, particularly by the type of comparison groups, assumptions about lead time, and the denominator used to calculate the rates. The different methodologies led to wide variations in estimates and a lack of agreement as to the true rate of overdiagnosis from mammography screening.

New Studies

A meta-analysis of three RCTs, a systematic review of 13 observational studies, and 17 individual studies of overdiagnosis were identified for the current update (Table 20). Estimates of overdiagnosis were primarily based on screening trials, screening programs and registries, or modeled data. Studies differed by patient populations of various ages and with different risks for breast cancer; screening and followup times; screening policies, uptake, and intensity; and underlying cancer incidence trends. Estimates differed in their numerators and denominators; whether they included both invasive cancer and DCIS; 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 differences between the incidence of cancer in the presence and in the 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). In addition, at least seven different measures of overdiagnosis were reported in published papers. How differences in study characteristics, methods, and measures effect estimates of overdiagnosis have been well described, yet there is currently no consensus about the most appropriate approach.

Estimates from RCTs. Data from three RCTs that did not provide screening of controls at the
end of the trial were considered to be the least biased estimates of overdiagnosis in a comprehensive review commissioned by Cancer Research U.K. and the Department of Health in England. The Malmö I and Canadian National Breast Screening Study (CNBSS-1 and CNBSS-2) trials provided randomized comparison groups, and followup times that extended sufficiently beyond the screening period to differentiate earlier diagnosis from overdiagnosis.

Using results of the Malmö I and two Canadian trials, the excess incidence of breast cancer (both invasive cancer and DCIS) in the screening population was compared with the incidence in the absence of screening (Table 21). For the short case accrual method that includes cases identified only during the screening period, overdiagnosis was estimated to be 19.0 percent (95% CI, 15.2% to 22.7%; \(I^2=64.8\%; p=0.058; 3\) trials). For the long case accrual method that includes cases identified throughout the screening and followup periods, overdiagnosis was 10.7 percent (95% CI, 9.3% to 12.2%; \(I^2=22.3\%; p=0.276; 3\) trials). Estimates for women age 40 to 49 years in the CNBSS-1 trial were higher (22.7% for short case accrual; 12.4% for long case accrual) than for women age 50 to 59 years in the CNBSS-2 trial (16.0% and 9.7%, respectively), and women age 55 to 69 years in the Malmö trial (18.7% and 10.5%, respectively).

However, overdiagnosis estimates from the trials included in this meta-analysis used different denominators. The Malmö I trial included all breast cancer cases, not just those identified with screening, while the Canadian trials included only cancer cases detected by screening. If these were calculated similarly, results from Malmö would be 23 percent instead of the 11 percent estimate used. In addition, more recently published long-term followup of the two Canadian trials 15 years after enrollment indicated a 22 percent overdiagnosis rate for combined age groups.

Data from the other RCTs are susceptible to over- or underestimating overdiagnosis because women in the control groups were offered screening at the end of the screening periods. If cases from screened control groups were included in the estimate of overdiagnosis, differences between comparison groups would be reduced and overdiagnosis would be underestimated. If these cases were excluded, overdiagnosis estimates would be inflated because the control group would not have been followed up long enough to determine accurate estimates. New publications of overdiagnosis reported from trials that screened control groups indicated none or minimal overdiagnosis, including the Swedish Two-County trial and Screening for Young Women Trial.

** Estimates from screening programs and registries.** A systematic review for the EUROSCREEN Working Group included 13 observational studies providing estimates of overdiagnosis in European population-based screening programs. Five newer studies in this review are included in this update, and three older studies were included in the previous report (Table 20). These studies differed by many of the study characteristics, methods, and measures previously described that limited comparisons and prohibited combined estimates. In particular, for studies comparing screening and nonscreening populations from different time periods, adjustments for breast cancer risk were dependent on correct estimates of temporal trends. Also, denominators defining the populations at risk were inconsistent across studies (e.g., breast cancer diagnosis in an entire population versus women of a specific age who attended screening). Importantly, most studies used denominators that included all breast cancer
cases, rather than screen-detected cases, leading to lower estimates.

Unadjusted estimates from the 13 observational studies included in the EUROSCREEN review indicated overdiagnosis rates ranging from 0 to 54 percent. For six studies that adjusted overdiagnosis estimates for breast cancer risk and lead time, rates varied from 1 to 10 percent.

European studies published since the EUROSCREEN systematic review include three studies of the Norwegian Breast Cancer Screening Program (NBCSP), and one from Denmark. The Norwegian studies reported overdiagnosis rates of 13.9 to 16.5 percent of invasive cancer + DCIS, and 9.6 to 11.3 percent of invasive cancer only, when comparing women screened with those never invited or nonattenders; 15 to 25 percent of invasive cancer depending on region and lead time assumptions when comparing populations in regions with versus without screening; and 50 percent of invasive cancer + DCIS when comparing screening versus discontinuation of screening that assumes that all increases in incidence were due to overdiagnosis. The Danish study estimated overdiagnosis as the cumulative incidence of breast cancer + DCIS in regions with screening compared with expected cumulative incidence. For women followed for at least 8 years, the estimates were 3 percent in Copenhagen and 0.7 percent in Funen. These results contrast with the estimate of 33 percent overdiagnosis in an earlier study of the same regions that used ratios of incidence between screened and nonscreened areas for the screened age group.

An analysis of the Canadian British Columbia Cancer Registry between 1970 and 2009 provided two estimates of overdiagnosis in women age 40 to 89 years. Rates were 17.3 percent of invasive cancer + DCIS and 5.4 percent for invasive cancer only when using cumulative incidence rates of women involved in active screening compared with women who were never screened or not actively screened. A second estimate compared the observed and expected cumulative population incidence rates between two time periods, resulting in estimates of 6.7 percent of invasive cancer + DCIS and -0.7 percent for invasive cancer only.

The only study conducted in the United States was not based on screening programs, but on comparisons of the expected increase in the incidence of early-stage cancer detected with mammography screening with the actual decrease in late-stage cancer incidence in women over a 40-year period using Surveillance, Epidemiology, and End Results (SEER) data. Overdiagnosis rates based on different incidence trend assumptions were estimated from 22 to 31 percent.

Estimates from models. Although several statistical models of overdiagnosis have been published, these studies have been less acceptable to guideline development groups because of the many assumptions that were used to construct them. Results of the models are heavily dependent on estimates of lead time and of progression rates from DCIS to invasive cancer. The longer the lead time, the more the estimate decreases with time. Assignment of these assumptions can be subjective. Six new studies of models of screening populations in The Netherlands, United Kingdom, France, Spain, and Australia, met inclusion criteria for this update (Table 20).

A microsimulation model for estimating overdiagnosis in screening programs in The Netherlands
provided 1-year estimates across different time periods that ranged from 1.0 to 11.4 percent.\textsuperscript{168} Another microsimulation model of screening programs in The Netherlands provided estimates based on assumptions of the progression from DCIS to invasive cancer.\textsuperscript{167} These included overdiagnosis estimates of 1.4 to 7.7 percent of all breast cancer, and 5.0 to 25.2 percent of screen-detected breast cancer.

A Markov simulation model estimated overdiagnosis for different screening strategies in the United Kingdom, including annual, triennial, and combination strategies for different age groups.\textsuperscript{180} For all invasive + DCIS cases diagnosed from age 40 to 85 years, overdiagnosis estimates ranged from 4.3 percent for triennial screening for women age 50 to 70 years, to 8.9 percent for annual screening from age 40 to 73 years. For screen-detected invasive + DCIS cases, overdiagnosis estimates ranged from 11.8 percent for triennial screening for women age 50 to 70 years, to 13.5 percent for annual screening from age 40 to 46 years followed by triennial screening from age 47 to 73 years.

A French study utilized a stochastic process for modeling all-cause mortality, lifetime probability of breast cancer, the natural course of breast cancer, and the detection of breast cancer clinically or by screening mammography.\textsuperscript{177} Overdiagnosis estimates included 1.5 percent of all diagnosed and 3.3 percent of screen-detected invasive cancer cases, and 28 percent of all diagnosed and 31.9 percent of screen-detected DCIS cases. In a Spanish population, a Poisson regression model was used to estimate expected incidence, accounting for age at diagnosis, reproductive factors, use of mammography, and year of birth.\textsuperscript{173} Estimates of overdiagnosis of invasive cancer varied from 0.4 percent in the oldest to 46.6 percent for the youngest cohort.

An Australian study estimated incidence in unscreened age groups (≤40 or ≥80 years) and in all age groups prior to implementation of screening adjusting for risk factors and lead time.\textsuperscript{174} Assuming a 2-year lead time, estimates from the first approach ranged from 27 to 66 percent, while estimates from the second approach were 36 to 47 percent. In general, rates were higher among women age 50 to 59 than 60 to 69.

Anxiety, Distress, and Other Psychological Responses

Previous Reports

A systematic review of 54 studies evaluated the adverse psychological effects of mammography screening programs.\textsuperscript{183} Most were cohort studies, and 24 used validated psychological measurement scales to assess the effects of screening. Studies indicated that women who received clear communication of their negative mammography results had minimal anxiety.\textsuperscript{183} Results were mixed in studies of women who were recalled for further testing as a result of screening. In several studies, women had persistent anxiety, despite eventual negative results, whereas some showed only transient anxiety.\textsuperscript{183} Some studies showed no differences between anxiety levels of women who had initial negative screening mammography results and those who had false-positive results.\textsuperscript{183}

A second systematic review of 23 studies (in 27 publications, of which 15 were included in the systematic review described above) specifically examined the effects of false-positive screening
mammography results on women age 40 years or older.\textsuperscript{184} Twenty studies were included that measured psychological distress, anxiety, and worry. False-positive mammography results had no consistent effect on most women’s general anxiety and depression but increased breast cancer-specific distress, anxiety, apprehension, and perceived breast cancer risk for some.\textsuperscript{184}

**New Studies**

A good-quality review of seven studies examined the effects of false-positive screening mammography results on women (Table 22).\textsuperscript{185} Three studies that evaluated breast cancer specific worry or distress reported significantly more distress among women with false-positive results, even after 35 months (1 study), than women with normal screening results. The most distress was observed among women who had biopsies (RR 2.07; 95% CI, 1.22 to 3.52), FNA (RR 1.80; 95% CI, 1.17 to 2.77), and early recall (RR 1.82; 95% CI, 1.22 to 2.72).\textsuperscript{183,186,187} Two studies that evaluated general anxiety and depression found no differences between women with true versus false-positive results.\textsuperscript{188,189} Among six studies that evaluated re-attendance rates after receiving a false-positive result, two studies reported that women with false-positive results were less likely to return for their next mammogram (RR 0.97; 95% CI, 0.96 to 0.98 and RR 0.92; 95% CI, 0.86 to 0.98);\textsuperscript{186,190} while two studies reported no differences.\textsuperscript{191,192} One study reported an increase in re-attendance when women were given letters tailored to their last screening result (RR 1.10; 95% CI, 1.00 to 1.21).\textsuperscript{193}

Another review rated fair-quality evaluated 17 studies of women age 40 to 74 years and reported that those with false-positive results had more anxiety, psychological distress, and breast cancer specific worry after screening (15 studies) compared with those with normal screening results.\textsuperscript{194} In two studies anxiety increased when women were recalled for biopsies,\textsuperscript{195,196} and in one study, anxiety persisted for 2 years after screening.\textsuperscript{197} The findings in these reviews are consistent with those from the previous report.\textsuperscript{183,184}

In addition to the reviews, 10 observational studies published after the reviews met inclusion criteria (Table 23). These include two fair-quality prospective cohort studies\textsuperscript{198,199} and three fair-quality retrospective cohort studies (Appendix B2-4);\textsuperscript{200-202} two good-quality\textsuperscript{203,204} and one fair-quality\textsuperscript{205} nested case-control studies; one fair-quality case-control study;\textsuperscript{206} and one before-after study\textsuperscript{207} that was not quality rated because rating criteria are not available for this study design.

Five studies compared women receiving false-positive results with those receiving normal screening results,\textsuperscript{198,203-206} and reported similar findings as the reviews. Women with false-positive versus normal screening results experienced more breast cancer worry (49\% vs. 10\%, p<0.0001) and had more worries that affected mood or daily activities (31\% vs. 2\%, p<0.0001).\textsuperscript{206} These women also had lower mental functioning and vitality measured by the Short Form 36 Health Survey (SF-36) at 6 months (mean mental functioning score: 80.6 vs. 85.0; p=0.03; mean vitality score: 70.3 vs. 77.0; p=0.02).\textsuperscript{205} A study of 323 participants reported higher depression scores on the Hospital Anxiety and Depression Scale (HADS) – Depression Subscale (HADS-D) for women with false-positive versus normal screening results (6-months mean: 3.2 vs. 2.4, p=0.045), however neither group reached clinical thresholds.\textsuperscript{205} Other studies of general anxiety and depression measured with the HADS or State-Trait Anxiety Inventory (STAI) reported no significant differences between groups.\textsuperscript{198,204,206} However, in a study of
13,491 women, analysis of racial sub-groups indicated increased depression among non-white women (6%; n=847) with false-positive results (OR 3.23; 95% CI, 1.32 to 7.91).198

Psychological outcomes of women with false-positive, normal screen, and true-positive results (breast cancer diagnosis) were compared in a good-quality nested case-control study using the Consequences of Screening in Breast Cancer questionnaire.203 Immediately after screening, women with normal screening results had better scores on all subscales compared with women with either false-positive or true-positive results (p<0.001 for all outcomes), but there were no differences between women with false-positive and true positive results. By 3 years after screening, women with normal screening results continued to have better scores on all subscales compared with women with true-positive results (p<0.002 for all outcomes). However, women with normal screening results also had better scores than those with false-positive results on subscales for sense of dejection, anxiety, negative impact on behavior or sleep, social network, existential values, and on single items of feeling less attractive and keeping mind off things (p<0.03 for all outcomes). Women with false-positive results had better scores than women with true-positive results on all but the breast examination and worried about breast cancer subscales (p<0.03 for all outcomes).

Anxiety and depression were evaluated in a before-after study of women with false-positive and true-positive results at the time of mammography recall and 4 weeks after.207 The proportion meeting the HADS threshold for anxiety (score >11) decreased from recall to 4 weeks for women with false-positive (15% to 5.5%) and true positive results (19% to 17%). The proportion meeting the HADS threshold for depression (score >11) also decreased from recall to 4 weeks (1.4% to 1.3%) for women with false-positive results, but increased for women with true positive results (1.3% vs. 6.9%). In multivariate models, factors predicting anxiety or depression at followup included low general life expectations, previous history of anxiety and/or depression, and anxiety at baseline. Satisfaction with information also predicted depression. Anxiety and depression were also evaluated in a fair-quality prospective cohort study of 482 women that compared women recalled after their first screening mammography with those recalled after a repeat screening mammography.199 Both groups had similar anxiety and depression scores initially that significantly declined over the following 6 months.

Three studies compared re-attendance rates of women with false-positive screening mammography results with women with normal screening results.200-202 As with the systematic review of studies on re-attendance, the results of these studies were also inconsistent. One study reported higher re-attendance rates for women with normal results (93.2% vs. 52.1% for false-positive result), and the lowest rates for women recalled to screening more than once for different lesions (44.3%).202 The other two studies reported higher rates of re-attendance for women with false-positive compared with normal screening results (90.7% vs. 89.0%, p<0.001200 and 87.7% vs. 86.0%, difference of 1.61%, 95% CI, 0.54% to 2.62%).201 The OR for re-attendance was higher for women who did not receive tissue sampling after false-positive versus normal mammography screening (OR 1.20, 95% CI, 1.10 to 1.30).201 Older women had lower odds of re-attendance at both prevalent (OR 0.89, 95% CI 0.86 to 0.93) and incident screening rounds (OR 0.99, 95% CI 0.98 to 0.99).

In a study of women with false-positive results, ORs for re-attendance were lower for women
receiving open biopsies (adjusted OR [AOR] 0.4, 95% CI, 0.3 to 0.6) but not core needle biopsies compared with women receiving no tissue sampling. This study also found that older women had reduced odds of re-attendance (AOR 0.8, 95% CI, 0.7 to 0.9 for women aged 55 to 59 years and 0.8, 95% CI, 0.6 to 0.9 for women aged 60 to 62 years compared with women aged 50 to 54 years).

Radiation Exposure

Previous Reports

In the previous report, estimates of radiation exposure were provided by a systematic review that included various types of studies of radiation exposure as a basis for predicting risk for inducing breast cancer. However, these estimates were not specific for radiation induced risk or mortality attributable to mammography or breast imaging.

New Studies

No studies directly measured the association between radiation exposure from mammography screening and the incidence of breast cancer and death for film, digital, or tomosynthesis. The general concern about the harms of radiation exposure stems from the assumption that higher doses of radiation induce cancers. Two-view digital mammography and screen-film mammography involve an average mean glandular radiation dose (MGD) of 3.7 and 4.7 mGy, respectively, and are considered low dose, low energy radiation. Radiation exposure for tomosynthesis is generally considered to be up to twice the dose of digital mammography.

Two modeling studies provided estimates of radiation exposure, breast cancer incidence, and death (Table 24). In a study based on theoretical estimates, the average estimated MGD and the lifetime attributable risk (LAR) of radiation induced breast cancer incidence and mortality were calculated based on age-specific estimates in the United States screening population. Results indicated that a 40 year old woman undergoing a single, bilateral, two-view screening mammogram has an LAR of breast cancer incidence of 5 to 7 cases per 100,000, and an LAR of breast cancer mortality of 1.3 to 1.7 deaths per 100,000. There was little effect on estimated risk when screening ended at age 80 years or later. Risks were similar for digital breast tomosynthesis (LAR 1.3 to 2.6 deaths).

A modeling study based on assumptions from the first study created an excess absolute risk model to predict the number of radiation induced breast cancers attributable to the radiation dose received for a single typical digital mammogram. Results indicated that the estimated number of deaths due to radiation induced cancer was between 2 per 100,000 in women age 50 to 59 years screened biennially, and up to 11 per 100,000 in women screened annually between ages 40 to 59 years. Women age 40 to 49 years undergoing annual mammographic screening would have an absolute risk of radiation induced mortality of 7.6 per 100,000. The calculations in this study are based on radiation doses from digital mammography, whereas previous estimates based on film mammography used higher doses per examination (3.7 mGy vs. 4.5 mGy, respectively).
Pain During Procedures

Previous Reports

A systematic review of 22 studies of pain and discomfort associated with mammography indicated that many women experience pain during the procedure (range, 1% to 77%), but few would consider this a deterrent from future screening. In these studies, pain was associated with the stage of the menstrual cycle, anxiety, and the anticipation of pain.

A good-quality systematic review of seven intervention trials to reduce pain with screening mammography indicated that discomfort was reduced when written or verbal information was provided to women, and when a breast cushion was used. Use of different breast compression strategies or premedication with acetaminophen had no significant effects in reducing discomfort.

New Studies

Breast compression is used during mammography to create uniform density, reduce breast thickness, and flatten overlying skin and tissues, which contributes to sharper images and reduces the radiation dose. However, compression may add to the discomfort of mammography for some women.

A good-quality recent review of 20 observational studies, most cross-sectional, examined pain or discomfort after screening mammography and its effect on re-attendance for future screening mammography (Table 25). Seven studies reported the proportion of women who experienced pain with previous mammography who directly stated this as their reason for non-re-attendance. In these studies, actual non-re-attendance indicating pain as the reason ranged from 11 to 46 percent (5 studies), and intended future non-re-attendance because of pain ranged from 3 to 18 percent (2 studies).

Fifteen studies reported the proportion of women who experienced pain with previous mammography and the proportion of women who re-attended as an outcome, but did not directly ask non-re-attendees for their reasons. There was no difference in actual re-attendance between women who experienced pain and those who did not (RR 1.38; 95% CI, 0.94 to 2.02; 5 studies). However, non-re-attendees had significantly higher pain scores compared with re-attendees in two of three studies. Two studies reported less intent to re-attend for women with pain, with OR 0.61 (95% CI, 0.38 to 0.98) in one study; while three others reported no differences in intended re-attendance and pain. This review is consistent with findings from the previous report.
Key Question 6. How Do the Harms of Routine Breast Cancer Screening Vary by Different Screening Modality?

Summary

• Four of five observational studies demonstrated statistically significantly lower rates of recall for tomosynthesis and mammography compared with mammography alone.
• A U.S. study comparing tomosynthesis and mammography with mammography alone reported a reduction of 16 recalls per 1,000 women and an increase in cancer detection of 1.2 cases per 1,000 women, but also an increase of 1.3 biopsies per 1,000 women. Another U.S. study reported a 38 percent reduction in recall rates when tomosynthesis was added to digital mammography versus mammography alone.
• Women receiving mammography and CBE compared with mammography alone had higher recalls in a study from Canada (55 per 10,000 additional recalls with CBE).
• No studies evaluated screening with ultrasound or MRI in women who are not at high risk for breast cancer.

Evidence

Previous Reports

Previous evidence reviews for the USPSTF did not address this question.

New Studies

There are no RCTs of screening using tomosynthesis, ultrasound, or MRI in women who are not at high risk for breast cancer. Six observational studies compared false-positive recall rates of screening for breast cancer using mammography and tomosynthesis,146,147,212-214 or CBE215 compared with mammography alone (Table 26). No studies evaluated MRI screening in women who are not at high-risk for breast cancer. Use of supplemental imaging for women with dense breasts is included in a separate report.

Four of five studies demonstrated statistically significantly lower rates of recall for tomosynthesis and mammography compared with mammography alone.146,147,212-214 One of the U.S. studies reported a reduction of 16 recalls per 1,000 women (95% CI, -18 to -14, p<0.001), increase of 1.3 biopsies per 1,000 women (95% CI, 0.4 to 2.1; p=0.004), increase in cancer detection of 1.2 per 1,000 women (95% CI, 0.8 to 1.6; p<0.001), and increase in invasive cancer detection of 1.2 per 1,000 women (95% CI, 0.8 to 1.6; p<0.001).212

Recall reductions were not statistically significant in another smaller U.S. study.146 Importantly, there was an overall reduction in false positives and an increase in biopsies, accompanied by an increase in cancer detection involving only invasive cancers, regardless of breast density or age.212 Another smaller, U.S. observational study demonstrated reduced recall rates with tomosynthesis after controlling for age, breast density, and breast cancer risk (AOR 0.62, 95%
CI, 0.55 to 0.70; p<0.0001) versus mammography alone. Two European studies also found significantly lower rates of recalls for women screened with tomosynthesis and mammography (1% vs. 2%, p<0.0001; and 53/1,000 vs. 61/1,000; p=0.001). Women receiving mammography and CBE compared with mammography alone had higher recalls in a study from Canada (8.7% vs. 6.5%; 55/10,000 additional recalls with CBE).

Contextual Question 1. What Are the Rates of Specific Adverse Effects of Current Treatment Regimens for Invasive Breast Cancer and DCIS in the United States?

Rates of specific adverse effects of breast cancer treatment regimens are not provided in centralized sources, but rather the available information is found in publications of surgical case series, clinical trials, and information from drug package inserts. Examples of rates of several recommended and commonly used treatments in the U.S. are summarized in Table 27.

Most patients with DCIS and Stage I to III invasive cancer receive surgery, including lumpectomy or mastectomy with sentinel lymph node biopsy and, with more extensive disease, axillary lymph node dissection. Many will also undergo reconstruction surgery. The most common adverse effects include wound infection, skin flap necrosis, and chronic chest wall pain. Approximately 5 percent of patients with sentinel lymph node biopsy and 16 to 18 percent with axillary lymph node dissection develop clinical lymphedema. Some patients experience phantom breast syndrome, pneumothorax, and brachial plexopathy.

Radiation therapy is provided to women with DCIS and Stage I to III disease, and with increasing frequency as the stage of disease progresses. Adverse effects to radiation therapy vary by dose and regimen. For example, among women with breast conserving surgery, a dose of 50 Gy (unit of radiation) in 25 fractions over 5 weeks may cause breast shrinkage in 25 percent, breast induration 18 percent, telangiectasia 5 percent, and breast edema 10 percent. Symptomatic rib fracture, lung fibrosis, ischemic heart disease and brachial plexopathy occur in less than 5 percent.

Endocrine therapy for 5 to 10 years, depending on the drug, is indicated for patients with ER positive DCIS and Stage I to III disease. For some women, tamoxifen causes hot flashes, vaginal discharge, and irregular menses. Less common adverse effects are thromboembolism, endometrial cancer, and cataracts. Anastrozole or other aromatase inhibitors are alternatives to tamoxifen that may cause hot flashes and joint pain. Less common adverse effects are vaginal bleeding, vaginal discharge, thromboembolic events, cataracts, and carpal tunnel syndrome.

Several neoadjuvant/adjuvant chemotherapy regimens are available to treat patients with Stage I to III disease, and selection is based on ER, PR, and HER2 status. Adverse effects include short-term (hair loss, nausea, vomiting, fatigue, neuropathy, neutropenia) and long-term (persistent neuropathy, heart failure) adverse effects that depend on regimen, duration, and age (examples in Table 27). Chemotherapy regimens for Stage IV disease are usually provided over extended periods of time because Stage IV disease is not curable. While extended treatment regimens can
control the disease for variable amounts of time depending on disease biology, they may have many adverse effects. These include neutropenia, fatigue, anemia, neuropathy, hair loss, nausea, and stomatitis, among others.

**Contextual Question 2. What Are the Absolute Incidence Rates of DCIS and Localized and Advanced Invasive Breast Cancer in Screened and Nonscreened Populations in the United States?**

Absolute incidence rates for DCIS and localized and advanced invasive breast cancer are not provided according to screened and nonscreened populations in the United States. The majority of cases of DCIS are identified by mammography screening and the increased incidence of DCIS corresponds to the advent of widespread screening.32 The most recent rates from SEER for invasive cancer include 129.6 per 100,000 for all age groups; 45.2 per 100,000 for age less than 50 years; and 350.4 per 100,000 for age 50 years or greater.220 Rates of DCIS include 35.5 per 100,000 for all age groups; 14.4 per 100,000 for less than 50 years, and 100.0 per 100,000 for age 50 years or greater.220

**Contextual Question 3. How Do Women Weigh Harms and Benefits of Screening Mammography, and How Do They Use This Information in Their Decisions to Undergo Screening?**

Research that describes how women weigh the benefits and harms of screening mammography and use this information for clinical decision making is limited. A Cochrane review of RCTs evaluating the effects of personalized risk communication on informed decision making found no association between provision of numerical information and uptake of mammography for women 40 years or older (OR 0.95; 95% CI, 0.78 to 1.15; 6 trials).221 However, there was an association for greater uptake of mammography when categorical information was given compared with general risk information (OR 1.29; 95% CI, 1.11 to 1.51; 6 trials). The review found that 45 percent (592/1309) of those who received personalized risk information made informed choices, compared with only 20 percent (229/1135) of those who received generic risk information (OR 4.48; 95% CI, 3.62 to 5.53; 3 trials).

Four main themes describing factors that influence a woman’s decision to attend breast cancer screening were identified in a review of 12 observational studies.222 These included psychological and practical factors; issues related to ethnicity; influence of socioeconomic status; and issues related to the screening program. In these studies, cancer anxiety and worry was associated with both the promotion and avoidance of breast cancer screening.223,224 Some women cited embarrassment as their reason for non-attendance,224 particularly women of specific religious groups.225 Most women expressed a preference for a female medical professional performing the screening mammography.224 Black women were more likely to get information about mammography from their primary physician, while white women were more likely to have received their information from media sources.226
In these studies, rates of screening uptake were lower among low-income populations and non-English speakers, and higher income households were twice as likely to attend mammography screening. Lower uptake rates were also associated with lower levels of education, the lack of health insurance, and unemployment. Women from lower socioeconomic backgrounds did not consider themselves at risk for breast cancer and focused on perceived negative aspects of screening and the intrinsic costs (time, embarrassment, and discomfort). In contrast, many women overestimated their risk and the mortality reduction from mammography screening resulting in higher uptake of screening.
Chapter 4. Discussion

Summary of Review Findings

Table 28 summarizes the evidence reviewed for this update and Table 29 provides a concise summary of benefits and harms. Trials of mammography screening indicated reduced breast cancer mortality with screening for women age 39 to 69 years, although results for ages 39 to 49 and 50 to 59 years were of borderline statistical significance and varied depending on how cases were accrued in trials. The absolute breast cancer mortality reduction per 10,000 women screened for 10 years varied from 4 for age 39 to 49 years; 5 to 8 for age 50 to 59 years; and 12 to 21 for age 60 to 69 years. Estimates for age 70 to 74 years were limited by low numbers of events in trials that had smaller numbers of women in this age group. The meta-analysis results reflect updated data from the Canadian (CNBSS-1 and CNBSS-2), Swedish Two-County Study, and Age trials that were not available for the previous review, as well as previously published results from the Stockholm, Gothenburg, Malmö (MMST I and MMST II), and HIP trials. The meta-analyses used long and short case accrual methods in order to explore the methodological differences of the trials and interpret findings using both approaches.

Observational studies of population-based mammography screening reported a wide range of reductions in breast cancer mortality. Most studies were conducted in Europe or the United Kingdom and included women age 50 to 69 years. Meta-analyses of studies indicated a breast cancer mortality RR of 0.75 (95% CI, 0.69 to 0.81) based on seven incidence-based mortality studies; and an OR of 0.69 (95% CI, 0.57 to 0.83) based on seven case-control studies. The 25 to 31 percent mortality reduction from observational studies compares with a 19 to 22 percent reduction estimated from the meta-analysis of screening trials for women age 50 to 69 years. A large observational study of Swedish women in their 40s indicated 26 percent reduction in breast cancer mortality for women invited to screening, while a Canadian study indicated 44 percent reduction for screening participants. These mortality reductions compare with 12 to 16 percent mortality reductions in the trials, although the trial estimates were only of borderline statistical significance.

All-cause mortality did not differ between randomized groups in meta-analyses of trials, regardless of whether trials were analyzed in combined or separate age groups. Also, no trials evaluated mortality outcomes on the basis of risk factors besides age, and there are no head-to-head trials of the effectiveness of different screening intervals or modalities.

The screening trials also provided several measures of intermediate breast cancer outcomes. When thresholds for advanced disease were defined by the most severe categories available from the trials (Stage III + IV disease, size ≥50 mm), a meta-analysis indicated a significant reduction in advanced disease for women age 50 years and older randomized to screening versus nonscreening groups (RR 0.62; 95% CI, 0.46 to 0.83; 3 trials), but not for women age 39 to 49 years. This reduction in intermediate outcomes aligns with the reduction in mortality outcomes that were also statistically significant in the trials for the older age groups.

Although no trials evaluated the incidence of advanced breast cancer outcomes and treatment on
the basis of risk factors or screening intervals, an analysis of BCSC data indicated a lower proportion of Stage III + IV disease among women age 40 to 49 years screened annually versus biennially. Also, BCSC data indicated that women age 40 to 49 years with extremely dense breasts had increased risks for advanced stage cancer (IIB+) and large-size tumors (>20 mm) with biennial compared with annual screening. These results suggest that women in their 40s with increased risk may reduce their risk for higher stage tumors with screening, even though mortality outcomes were not significantly reduced in the trials. These findings are consistent with a modeling study based on BCSC data that indicated that women in their 40s with 2-fold increases in risk (such as with extremely dense breasts) would experience benefits and harms comparable with average-risk women in their 50s when using life-years as the benefit metric.66

A Cochrane review that included five screening trials indicated that women randomized to screening were significantly more likely to have surgical and radiation therapy, and less likely to have hormone therapy than controls, while use of chemotherapy was similar between groups. This finding would be expected because screening increases detection of DCIS and early stage disease that are currently aggressively treated. However, treatment outcomes in the RCTs represent outdated therapies that limit their applicability. Observational studies of the impact of screening on advanced cancer diagnosis and treatment generally provided comparisons between screen-detected and nonscreened-detected cases rather than rates between screening populations that more directly address this Key Question.

There were few studies meeting inclusion criteria that compared the effectiveness of screening across various modalities, despite the increasing use of them in clinical practice. Tumor size, stage, and node status did not differ between women screened with tomosynthesis and digital mammography compared with those receiving mammography alone in two case series studies.

Several potential harms were also addressed in this systematic review. Updated BCSC data on digital mammography indicated that false-positive rates and recommendations for additional imaging were highest among women aged 40 to 49 years and declined with age, while false-negative rates were low across all age groups. Rates of recommendations for biopsy did not differ between ages. Results did not differ by time since last mammography screening regardless of whether broad or narrow estimates of one versus two years were used.

Several risk factors were statistically significantly associated with higher rates of false-positive and false-negative results and recommendations for additional imaging and biopsy across most age groups. These included family history of breast cancer, high breast density, and previous benign breast biopsy. Premenopausal status, use of menopausal hormone therapy, and lower BMI were associated with some of the outcomes for specific age groups only. Comparisons based on race and ethnicity indicated the lowest rates of false-positive results and additional imaging among Asians. Comparisons based on different combinations of breast density categories indicated that rates for all outcomes were lowest for women with almost entirely fat breasts, and highest for women with heterogeneously dense breasts or for those in the combined category of heterogeneous and extreme density. Women with extremely dense breasts had the highest rates of false-negative results.

While some risk factors reflect high exposure to estrogen and related changes in breast tissue
(premenopause, menopausal hormone therapy), others may serve primarily as markers of increased breast cancer risk (family history, previous benign biopsy). The mechanisms of these risk factors and whether screening outcomes are influenced by how they affect the mammographic image, increase clinical suspicion, or other ways, are beyond the scope of this analysis.

Additional publications of BCSC data indicated that 10-year cumulative rates of false-positive mammography and biopsy results were higher for annual than biennial screening (mammography 61% vs. 41%; biopsy 7% vs. 5%); for women with heterogeneously dense or extremely dense breasts; women in their 40s; and those who used combination hormone therapy.

Studies of overdiagnosis were primarily based on screening trials, screening programs and registries, or modeled data. Studies differed by their characteristics, methods, and measures, and estimates of the magnitude of overdiagnosis varied depending on the analytic approach, particularly regarding the different denominators used in the estimates. These estimates are difficult to apply to individual women because it is not known which types of cancer will progress, how quickly cancer will advance, and expected lifetimes.

Estimates of overdiagnosis from three RCTs that did not provide screening of controls at the end of the trial (Malmö I, CNBSS-1, CNBSS-2) indicated overdiagnosis rates of 11 to 22 percent. Unadjusted estimates from 13 observational studies indicated rates ranging from 0 to 54 percent; while six studies that adjusted for breast cancer risk and lead time indicated rates ranging from 1 to 10 percent.

Women with false-positive results were more distressed than women with negative results. Anxiety improved over time for most women, but persisted for over 2 years for some. Some women with false-positive results were less likely to return for their next mammogram, although studies were inconsistent. Although many women experienced pain during mammography (1% to 77%), the proportion of those experiencing pain who did not attend future screening varied (11% to 46%). Trials of interventions indicated that discomfort was reduced by providing written or verbal information or using breast cushions.

A U.S. study comparing tomosynthesis and mammography with mammography alone reported a significant reduction of 16 recalls per 1,000 women, but also an increase of 1.3 biopsies per 1,000 women. Mammography and CBE resulted in 55 per 10,000 additional recalls. Studies of screening with MRI or ultrasound focus on high-risk women. The number of deaths due to radiation induced cancer from screening with digital mammography was estimated through modeling as between 2 to 11 per 100,000 depending on age at onset and screening intervals. However, these models are based on assumptions that may not be accurate.

**Limitations**

Limitations of this review include using only English-language articles, which could result in language bias, although we did not identify non–English-language studies that otherwise met inclusion criteria in our searches. We only included studies that are applicable to current practice
in the United States in order to improve clinical relevance for the USPSTF, excluding much research in the field. This perspective may not be as relevant to other populations and settings. Despite using updated data, the RCTs of screening represent older technologies and cancer treatments that are not relevant today. Also, this update prioritized studies that addressed the Key Questions guiding the review. As with most areas of medicine, breast cancer screening does not exist in isolation, and applying inclusion criteria for studies places artificial boundaries around this complex topic. Many important issues could not be addressed because they were outside the scope of this review, including additional benefits (e.g., increasing breast awareness) and harms (e.g., economic hardship). Studies were lacking for some Key Questions, and the number, quality, and applicability of studies varied widely.

Emerging Issues and Next Steps

Breast cancer is a continuum of entities, not just one disease, that must be considered when choosing screening and treatment options and when balancing benefits and harms. None of the screening trials consider breast cancer in this manner. As diagnostic and treatment experiences become more individualized and include patient preferences and decision making, it becomes even more difficult to characterize benefits and harms in a general way. Many patients would consider quality-of-life issues important outcomes, although these issues are more difficult to measure and report in research studies.

New technologies, such as tomosynthesis and MRI, are becoming more widely used in the United States without definitive studies of their effects on screening outcomes. Consumer expectations that new technology is better than old may obscure potential adverse effects, such as higher false-positive results, biopsies, and expense. No screening trials incorporating newer technologies have been published, and estimates of benefits and harms in this report are based predominantly on studies of film and digital mammography. No trials have evaluated the appropriate interval for mammography screening or the role of risk factors.

Relevance for Priority Populations

Women age 70 years and older are a rapidly growing population in the United States, yet research on breast cancer screening and prevention in this age group is limited. Observational studies suggest that older women may benefit from regular mammography screening.

Most of the screening trials and studies of screening programs were based in Europe and the United Kingdom, and enrolled predominantly white women. Data on race and ethnicity from the BCSC suggest possible differences between groups, but inferences from these subgroups are generally inconclusive because of lower numbers of participants and missing data.

Very little research has been conducted on women who are not screened in the United States, whether by choice, access, or other issues. Individuals who do not participate in screening and prevention services differ from those who do, and particularly differ from women who enroll in research studies. More information about this population could lead to improvements that could
serve them better than currently available services.

**Future Research**

Additional research on benefits and harms of mammography screening with quality-of-life outcomes, as well as morbidity and mortality outcomes, would provide further understanding of the implications of routine screening. Data for specific groups of women, based on risk, racial and ethnic background, access to screening, or existence of co-morbidities, for example, could inform screening practice. Studies of older women are essential in order to improve the evidence on screening for them including when to discontinue screening. Studies on the role of additional imaging modalities in screening are required in order to appropriately incorporate this technology in the screening process. More information on DCIS is needed, including its implications and outcomes. Distinguishing aggressive from non-aggressive forms of DCIS could lead to more selective treatment and reduce the consequences of overdiagnosis, particularly uncertainties regarding the transition from DCIS to invasive cancer, and lead time issues. Improving the methodology and assumptions involved in estimates of overdiagnosis would provide more meaningful understanding of this potential harm.

**Conclusions**

Trials indicate that mammography screening prevents 4 deaths per 10,000 women age 40 to 49 years after 10 years; 5 to 8 for age 50 to 59 years; and 12 to 21 for age 60 to 69 years; while estimates for age 70 to 74 years are limited by low numbers. These results are generally supported by observational studies of screening programs of women age 50 to 69 years. Higher stage tumors are also reduced with screening for women over age 50 years and for younger women with dense breasts who have annual compared with biennial screening. False-positive results and additional imaging are common, particularly for younger women and those with risk factors, while biopsies occur less often. Rates of false-negative results are low. Estimates of overdiagnosis based on trials ranged from 11 to 22 percent, while estimates based on observational studies ranged from 1 to 10 percent. Although RCTs are lacking, observational studies of tomosynthesis and digital mammography indicate reduced recalls, but increased cancer detection and biopsy rates. Mammography screening at any age is a tradeoff of a continuum of benefits and harms that varies on population and individual levels.
References


Abbreviation: MRI=magnetic resonance imaging.
Abbreviation: MRI=magnetic resonance imaging.
Figure 3. Analytic Framework and Key Questions

Key Questions:

In the target population of women age 40 years and 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║?
5. What are the harms¶ of routine mammography screening, and how do they differ by age, risk factor‡, and screening interval?
6. How do the harms¶ of routine breast cancer screening vary by different screening modality║?

Contextual Questions:

1. What are the rates of specific adverse effects of current treatment regimens for invasive breast cancer and ductal carcinoma in situ (DCIS) in the United States?
2. What are the absolute incidence rates of DCIS and localized and advanced invasive breast cancer in screened and nonscreened populations in the United States?
3. How do women weigh the harms and benefits of screening mammography and how do they use this information in their decisions to undergo screening?

*Excludes women with pre-existing breast cancer; clinically significant BRCA mutations, Li-Fraumeni syndrome, Cowden syndrome, hereditary diffuse gastric cancer, or other familial breast cancer syndromes; high-risk lesions (DCIS, LCIS, ADH, ALH); or previous large doses of chest radiation (≥ 20 Gy) before age 30.

†Addresses contextual question 1.

‡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 age >50 years, body mass index.

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

║Screening modalities include: mammography (digital, tomosynthesis), magnetic resonance imaging (MRI), ultrasound, and clinical breast examination (alone or in combination).

¶Harms include: false positive findings, anxiety, false positive biopsies, false negative findings, false reassurance, overdiagnosis and resulting overtreatment, and radiation exposure.

Abbreviations: ADH=atypical ductal hyperplasia; ALH=atypical lobular hyperplasia; BRCA=breast cancer gene; DCIS=ductal carcinoma in situ; Gy=gray (unit of absorbed radiation); KQ=key question; LCIS=lobular carcinoma in situ; MRI=magnetic resonance imaging.
Figure 4. Methods of Case Accrual in Trials

A
Screening intervention

B
Short case accrual

C
Long case accrual

D
Followup

Comparison of Accrual Methods

**Short case accrual (B)**
Includes deaths occurring among cases diagnosed during the screening intervention period (A), and in some trials, within an additional defined case accrual period.

- Includes fewer cases
- Reduces contamination of control group if screened during followup
- Can introduce bias in the absence of concurrent screening after the intervention

**Long case accrual (C)**
Includes deaths occurring among cases diagnosed during the screening intervention period plus the followup period (A+D).

- Includes more cases.
- Potential to dilute a true benefit because participants from the control group may be screened after the study intervention period has ended.
Figure 5. Meta-Analysis of Effects of Screening Trials on Breast Cancer Mortality With Longest Case Accrual Available

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study</th>
<th>Mean followup, year</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>30 to 49 years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nyström et al., 2002*</td>
<td>Malmö II</td>
<td>11.2</td>
<td>0.64 (0.39 to 1.06)</td>
</tr>
<tr>
<td>Tabár et al., 1995*</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.73 (0.37 to 1.41)</td>
</tr>
<tr>
<td>Tabár et al., 1995*</td>
<td>Östergötland</td>
<td>12.5</td>
<td>1.02 (0.52 to 1.99)</td>
</tr>
<tr>
<td>Moss et al., 2009*</td>
<td>Age</td>
<td>13.6</td>
<td>0.86 (0.71 to 1.04)</td>
</tr>
<tr>
<td>Bjurström et al., 2003*</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.69 (0.45 to 1.05)</td>
</tr>
<tr>
<td>Habbema et al., 1986*</td>
<td>HIP</td>
<td>14.0</td>
<td>0.75 (0.53 to 1.05)</td>
</tr>
<tr>
<td>Nyström et al., 2002*</td>
<td>Stockholm</td>
<td>14.3</td>
<td>1.52 (0.80 to 2.88)</td>
</tr>
<tr>
<td>Nyström et al., 2002*</td>
<td>Malmö I</td>
<td>18.2</td>
<td>0.74 (0.42 to 1.29)</td>
</tr>
<tr>
<td>Miller et al., 2014*</td>
<td>CNBSS-1</td>
<td>21.9</td>
<td>1.04 (0.87 to 1.24)</td>
</tr>
<tr>
<td><strong>Subtotal (P=25.4%, p=0.218)</strong></td>
<td></td>
<td></td>
<td>0.88 (0.73 to 1.003)</td>
</tr>
<tr>
<td>50 to 59 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tabár et al., 1995*</td>
<td>Östergötland</td>
<td>12.5</td>
<td>0.85 (0.52 to 1.38)</td>
</tr>
<tr>
<td>Tabár et al., 1995*</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.48 (0.29 to 0.77)</td>
</tr>
<tr>
<td>Nyström et al., 2002*</td>
<td>Stockholm</td>
<td>13.7</td>
<td>0.56 (0.32 to 0.97)</td>
</tr>
<tr>
<td>Bjurström et al., 2003*</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.83 (0.60 to 1.15)</td>
</tr>
<tr>
<td>Habbema et al., 1986*</td>
<td>HIP</td>
<td>14.0</td>
<td>0.83 (0.61 to 1.13)</td>
</tr>
<tr>
<td>Nyström et al., 2002*</td>
<td>Malmö I</td>
<td>18.1</td>
<td>0.98 (0.75 to 1.29)</td>
</tr>
<tr>
<td>Miller et al., 2014*</td>
<td>CNBSS-2</td>
<td>21.9</td>
<td>0.94 (0.78 to 1.13)</td>
</tr>
<tr>
<td><strong>Subtotal (P=38.0%, p=0.139)</strong></td>
<td></td>
<td></td>
<td>0.86 (0.68 to 0.97)</td>
</tr>
<tr>
<td>60 to 69 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tabár et al., 1995*</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.58 (0.35 to 0.96)</td>
</tr>
<tr>
<td>Tabár et al., 1995*</td>
<td>Östergötland</td>
<td>12.5</td>
<td>0.62 (0.43 to 0.91)</td>
</tr>
<tr>
<td>Nyström et al., 2002*</td>
<td>Stockholm</td>
<td>13.1</td>
<td>0.94 (0.46 to 2.02)</td>
</tr>
<tr>
<td>Habbema et al., 1986*</td>
<td>HIP</td>
<td>14.0</td>
<td>0.85 (0.48 to 1.47)</td>
</tr>
<tr>
<td>Nyström et al., 2002*</td>
<td>Malmö I</td>
<td>15.5</td>
<td>0.64 (0.45 to 0.92)</td>
</tr>
<tr>
<td><strong>Subtotal (P=0.0%, p=0.739)</strong></td>
<td></td>
<td></td>
<td>0.67 (0.54 to 0.83)</td>
</tr>
<tr>
<td>70 to 74 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tabár et al., 1995*</td>
<td>Östergötland</td>
<td>12.5</td>
<td>0.82 (0.43 to 1.58)</td>
</tr>
<tr>
<td>Tabár et al., 1995*</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.76 (0.42 to 1.36)</td>
</tr>
<tr>
<td>Nyström et al., 2002*</td>
<td>Malmö I</td>
<td>13.6</td>
<td>0.89 (0.15 to 6.60)</td>
</tr>
<tr>
<td><strong>Subtotal (P=0.0%, p=0.962)</strong></td>
<td></td>
<td></td>
<td>0.88 (0.31 to 1.28)</td>
</tr>
</tbody>
</table>

*Uses short case accrual, but these are the most inclusive results available.

**Abbreviations**: CI=confidence interval; CNBSS=Canadian National Breast Screening Study; HIP=Health Insurance Plan of Greater New York.
Figure 6. Meta-Analysis of Effects of Screening Trials on Breast Cancer Mortality With Short Case Accrual

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study</th>
<th>Mean followup, year</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>39 to 49 years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moss et al., 2008</td>
<td>Age</td>
<td>10.7</td>
<td>0.81 (0.64 to 1.02)</td>
</tr>
<tr>
<td>Nyström et al., 2002</td>
<td>Malmö II</td>
<td>11.2</td>
<td>0.64 (0.39 to 1.06)</td>
</tr>
<tr>
<td>Bjurman et al., 2003</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.56 (0.34 to 0.91)</td>
</tr>
<tr>
<td>Nyström et al., 2002</td>
<td>Stockholm</td>
<td>14.3</td>
<td>1.52 (0.60 to 2.88)</td>
</tr>
<tr>
<td>Shapiro et al., 1988</td>
<td>HIP</td>
<td>18</td>
<td>0.72 (0.49 to 1.06)</td>
</tr>
<tr>
<td>Nyström et al., 2002</td>
<td>Malmö I</td>
<td>18.2</td>
<td>0.74 (0.42 to 1.29)</td>
</tr>
<tr>
<td>Miller et al., 2014</td>
<td>CNBSS-1</td>
<td>21.9</td>
<td>1.09 (0.80 to 1.49)</td>
</tr>
<tr>
<td>Tabár et al., 2011</td>
<td>Two-county</td>
<td>25.7</td>
<td>0.95 (0.62 to 1.46)</td>
</tr>
<tr>
<td><strong>Subtotal (I²=35.8%, p=0.143)</strong></td>
<td></td>
<td></td>
<td>0.84 (0.70 to 1.00)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>50 to 59 years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nyström et al., 2002</td>
<td>Stockholm</td>
<td>13.7</td>
<td>0.56 (0.32 to 0.97)</td>
</tr>
<tr>
<td>Bjurman et al., 2003</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.93 (0.63 to 1.38)</td>
</tr>
<tr>
<td>Shapiro et al., 1988</td>
<td>HIP</td>
<td>18</td>
<td>0.80 (0.56 to 1.15)</td>
</tr>
<tr>
<td>Nyström et al., 2002</td>
<td>Malmö I</td>
<td>18.1</td>
<td>0.98 (0.75 to 1.29)</td>
</tr>
<tr>
<td>Miller et al., 2014</td>
<td>CNBSS-2</td>
<td>21.9</td>
<td>1.02 (0.77 to 1.36)</td>
</tr>
<tr>
<td>Tabár et al., 2011</td>
<td>Two-county</td>
<td>23.6</td>
<td>0.64 (0.45 to 0.91)</td>
</tr>
<tr>
<td><strong>Subtotal (I²=33.8%, p=0.182)</strong></td>
<td></td>
<td></td>
<td>0.85 (0.69 to 1.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>60 to 69 years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nyström et al., 2002</td>
<td>Stockholm</td>
<td>13.1</td>
<td>0.94 (0.46 to 2.02)</td>
</tr>
<tr>
<td>Nyström et al., 2002</td>
<td>Malmö I</td>
<td>15.5</td>
<td>0.64 (0.45 to 0.92)</td>
</tr>
<tr>
<td>Shapiro et al., 1988</td>
<td>HIP</td>
<td>18</td>
<td>0.97 (0.51 to 1.83)</td>
</tr>
<tr>
<td>Tabár et al., 2011</td>
<td>Two-county</td>
<td>18.6</td>
<td>0.61 (0.45 to 0.82)</td>
</tr>
<tr>
<td><strong>Subtotal (I²=0.0%, p=0.476)</strong></td>
<td></td>
<td></td>
<td>0.67 (0.55 to 0.91)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>70 to 74 years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tabár et al., 2011</td>
<td>Two-county</td>
<td>13.2</td>
<td>0.89 (0.57 to 1.39)</td>
</tr>
<tr>
<td>Nyström et al., 2002</td>
<td>Malmö I</td>
<td>13.6</td>
<td>0.98 (0.15 to 6.60)</td>
</tr>
<tr>
<td><strong>Subtotal (I²=0.0%, p=0.923)</strong></td>
<td></td>
<td></td>
<td>0.90 (0.46 to 1.78)</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI=confidence interval; CNBSS=Canadian National Breast Screening Study; HIP=Health Insurance Plan of Greater New York.
**Figure 7. Meta-Analysis of Effects of Screening Trials on All-Cause Mortality, With Combined Ages**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study</th>
<th>Age, year</th>
<th>Mean followup, year</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aron and Prorok, 1986**[^12^]</td>
<td>HIP</td>
<td>40 to 64</td>
<td>10</td>
<td>0.99 (0.93 to 1.05)</td>
</tr>
<tr>
<td>Nyström et al., 2002**[^7^]</td>
<td>Gothenburg</td>
<td>40 to 59</td>
<td>13.2</td>
<td>0.94 (0.88 to 1.00)</td>
</tr>
<tr>
<td>Nyström et al., 2002**[^7^]</td>
<td>Östergötland</td>
<td>40 to 74</td>
<td>17.2</td>
<td>0.98 (0.95 to 1.01)</td>
</tr>
<tr>
<td>Nyström et al., 2002**[^7^]</td>
<td>Stockholm</td>
<td>40 to 64</td>
<td>14.7</td>
<td>0.99 (0.95 to 1.03)</td>
</tr>
<tr>
<td>Nyström et al., 2002**[^7^]</td>
<td>Malmö II</td>
<td>43 to 49</td>
<td>9.1</td>
<td>1.03 (0.89 to 1.20)</td>
</tr>
<tr>
<td>Nyström et al., 2002**[^7^]</td>
<td>Malmö I</td>
<td>45 to 70</td>
<td>19.2</td>
<td>0.99 (0.97 to 1.01)</td>
</tr>
<tr>
<td>Moss et al., 2006[^9]</td>
<td>Age</td>
<td>40 to 49</td>
<td>10.7</td>
<td>0.97 (0.89 to 1.04)</td>
</tr>
<tr>
<td>Miller et al., 2014[^10]</td>
<td>CNBSS-1 &amp; 2</td>
<td>40 to 59</td>
<td>25</td>
<td>1.02 (0.98 to 1.06)</td>
</tr>
<tr>
<td>Overall (I² = 0.0%, p = 0.577)</td>
<td></td>
<td></td>
<td></td>
<td>0.99 (0.97 to 1.003)</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI = confidence interval; CNBSS = Canadian National Breast Screening Study; HIP = Health Insurance Plan of Greater New York.
Figure 8. Meta-Analysis of Effects of Screening Trials on All-Cause Mortality, Stratified by Age

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study</th>
<th>Mean followup, year</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 to 49 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tabár et al., 1989265</td>
<td>Kopparberg</td>
<td>7.9</td>
<td>1.33 (1.01 to 1.77)</td>
</tr>
<tr>
<td>Tabár et al., 1989265</td>
<td>Östergötland</td>
<td>7.9</td>
<td>0.93 (0.76 to 1.12)</td>
</tr>
<tr>
<td>Bjurstam et al., 1997242</td>
<td>Gothenburg</td>
<td>10.0</td>
<td>0.98 (0.86 to 1.12)</td>
</tr>
<tr>
<td>Frisell et al., 199780</td>
<td>Stockholm</td>
<td>11.0</td>
<td>1.12 (0.55 to 2.41)</td>
</tr>
<tr>
<td>Miller et al., 200276</td>
<td>CNBSS-1</td>
<td>13.0</td>
<td>1.00 (0.87 to 1.15)</td>
</tr>
<tr>
<td>Nyström et al., 200297</td>
<td>Malmö II</td>
<td>9.1</td>
<td>1.03 (0.89 to 1.20)</td>
</tr>
<tr>
<td>Moss et al., 200619</td>
<td>Age</td>
<td>10.7</td>
<td>0.97 (0.89 to 1.04)</td>
</tr>
<tr>
<td>Subtotal (F=0.0%, p=0.478)</td>
<td></td>
<td></td>
<td>0.99 (0.94 to 1.06)</td>
</tr>
</tbody>
</table>

| 50 to 59 years   |             |                     |                        |
| Tabár et al., 1989265 | Kopparberg | 7.9                | 1.00 (0.86 to 1.17)    |
| Tabár et al., 1989265 | Östergötland | 7.9                | 0.98 (0.87 to 1.11)    |
| Miller et al., 200070 | CNBSS-2   | 13.0               | 1.08 (0.96 to 1.20)    |
| Subtotal (F=0.0%, p=0.588) |     |                    | 1.02 (0.94 to 1.10)    |

| 60 to 69 years   |             |                     |                        |
| Tabár et al., 1989265 | Östergötland | 7.9                | 0.98 (0.91 to 1.05)    |
| Tabár et al., 1989265 | Kopparberg | 7.9                | 0.95 (0.87 to 1.04)    |
| Subtotal (F=0.0%, p=0.650) |     |                    | 0.97 (0.90 to 1.04)    |

| 70 to 74 years   |             |                     |                        |
| Tabár et al., 1989265 | Östergötland | 7.9                | 0.93 (0.87 to 1.01)    |
| Tabár et al., 1989265 | Kopparberg | 7.9                | 1.05 (0.95 to 1.15)    |
| Subtotal (F=72.4%, p=0.057) |     |                    | 0.98 (0.86 to 1.14)    |

Relative risk

Abbreviations: CI=confidence interval; CNBSS=Canadian National Breast Screening Study; HIP=Health Insurance Plan of Greater New York.
### Table 1: Meta-Analysis of Effects of Screening Trials on Advanced Cancer Outcome Using a Low Threshold for Advanced Cancer

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study</th>
<th>Definition of advanced cancer</th>
<th>Events, screening</th>
<th>Events, control</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>39 to 49 years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chu, et al., 1988&lt;sup&gt;11&lt;/sup&gt;</td>
<td>HIP</td>
<td>Stage II+</td>
<td>69/13,740</td>
<td>72/13,740</td>
<td>0.96 (0.89 to 1.33)</td>
</tr>
<tr>
<td>Tabár et al., 1995&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Swedish 2-county</td>
<td>Size ≥20 mm</td>
<td>85/19,844</td>
<td>75/15,604</td>
<td>0.89 (0.65 to 1.21)</td>
</tr>
<tr>
<td>Bjurström et al., 2003&lt;sup&gt;22&lt;/sup&gt;</td>
<td>Gothenburg</td>
<td>I+ lymph node</td>
<td>39/81,750</td>
<td>73/99,335</td>
<td>0.65 (0.44 to 0.96)</td>
</tr>
<tr>
<td>Autier et al., 2009&lt;sup&gt;26&lt;/sup&gt;</td>
<td>CNBSS-1</td>
<td>Size ≥ 20 mm</td>
<td>111/25,214</td>
<td>115/25,216</td>
<td>0.97 (0.74 to 1.25)</td>
</tr>
<tr>
<td>Autier et al., 2009&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Age</td>
<td>Size ≥ 20 mm</td>
<td>171/53,884</td>
<td>386/106,956</td>
<td>0.88 (0.73 to 1.05)</td>
</tr>
<tr>
<td><strong>Subtotal (I=0.0%, p=0.541)</strong></td>
<td></td>
<td></td>
<td>475/19,4432</td>
<td>721/260,851</td>
<td>0.88 (0.78 to 0.995)</td>
</tr>
<tr>
<td><strong>≥ 50 years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chu, et al., 1988&lt;sup&gt;11&lt;/sup&gt;</td>
<td>HIP</td>
<td>Stage II+</td>
<td>103/16,505</td>
<td>118/16,505</td>
<td>0.67 (0.67 to 1.14)</td>
</tr>
<tr>
<td>Tabár et al., 1995&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Swedish 2-county</td>
<td>Size ≥ 20 mm</td>
<td>389/57,236</td>
<td>432/40,381</td>
<td>0.64 (0.55 to 0.73)</td>
</tr>
<tr>
<td>Bjurström et al., 2003&lt;sup&gt;22&lt;/sup&gt;</td>
<td>Gothenburg</td>
<td>I+ lymph node</td>
<td>46/49,564</td>
<td>71/78,369</td>
<td>1.02 (0.71 to 1.48)</td>
</tr>
<tr>
<td>Autier et al., 2009&lt;sup&gt;26&lt;/sup&gt;</td>
<td>CNBSS-2</td>
<td>Size ≥ 20 mm</td>
<td>114/19,711</td>
<td>136/19,694</td>
<td>0.84 (0.65 to 1.07)</td>
</tr>
<tr>
<td><strong>Subtotal (I=70.4%, p=0.017)</strong></td>
<td></td>
<td></td>
<td>662/143,016</td>
<td>757/154,949</td>
<td>0.79 (0.64 to 1.03)</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI=confidence interval; CNBSS=Canadian National Breast Screening Study; HIP=Health Insurance Plan of Greater New York; mm=millimeter.

---

**Note:** The table presents a summary of meta-analysis results from various screening trials focusing on advanced cancer outcomes. The table includes the reference, study details, definitions of advanced cancer, event counts, and relative risks with 95% confidence intervals. The studies vary in their methodologies, including different thresholds for defining advanced cancer, and their outcomes are presented for different age groups. The relative risk values indicate the odds of advanced cancer outcomes in the screened groups compared to the control groups, with significance levels and confidence intervals provided for each study and subgroup analysis.
### Abbreviations

CI = confidence interval; CNBSS = Canadian National Breast Screening Study; HIP = Health Insurance Plan of Greater New York; mm = millimeter.
Table 1. Breast Cancer Staging System*

<table>
<thead>
<tr>
<th>Description</th>
<th>T1=tumor size ≤20 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary tumor (T)</td>
<td>T2=&gt;20 mm but ≤50 mm</td>
</tr>
<tr>
<td></td>
<td>T3=&gt;50 mm</td>
</tr>
<tr>
<td></td>
<td>T4=tumor of any size with direct extension to the chest wall and/or skin</td>
</tr>
<tr>
<td>Regional lymph nodes (N)</td>
<td>N0=no regional lymph node metastases</td>
</tr>
<tr>
<td></td>
<td>N1mi=micrometastases</td>
</tr>
<tr>
<td></td>
<td>N1=metastases to moveable ipsilateral axillary lymph nodes</td>
</tr>
<tr>
<td></td>
<td>N2=metastases in ipsilateral axillary lymph nodes that are clinically fixed</td>
</tr>
<tr>
<td></td>
<td>N3=metastases that are more extensive</td>
</tr>
<tr>
<td>Distant metastasis (M)</td>
<td>M0=no evidence of distant metastases</td>
</tr>
<tr>
<td></td>
<td>M1=distant detectable metastases as determined by clinical and radiographic means</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>DCIS</td>
</tr>
<tr>
<td>I</td>
<td>IA=T1, N0, M0</td>
</tr>
<tr>
<td></td>
<td>IB=T0, N1mi, M0 or T1, N1mi, M0</td>
</tr>
<tr>
<td>II</td>
<td>IIA=T0, N1, M0 or T1, N1, M0 or T2, N0, M0</td>
</tr>
<tr>
<td></td>
<td>IIB=T2, N1, M0 or T3, N0, M0</td>
</tr>
<tr>
<td>III</td>
<td>Larger size tumors with various combinations of lymph node involvement that are more extensive than stage II, but no distant metastases</td>
</tr>
<tr>
<td>IV</td>
<td>Distinct metastases (M1)</td>
</tr>
</tbody>
</table>

*Adapted from 2014 National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology.*

Abbreviations: DCIS=ductal carcinoma in situ; mm=millimeter.
Table 2. Treatment of DCIS and Invasive Breast Cancer by Stage*

<table>
<thead>
<tr>
<th>Treatment</th>
<th>0 (DCIS)</th>
<th>I, II, II, or T3, N1, M0</th>
<th>III (locally advanced)</th>
<th>IV (metastatic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgery</td>
<td>Total mastectomy ± sentinel node biopsy ± reconstruction; or lumpectomy</td>
<td>Total mastectomy or lumpectomy + axillary staging + breast reconstruction.</td>
<td>If response to pre-operative therapy, total mastectomy or lumpectomy + axillary dissection ± delayed breast reconstruction.</td>
<td>None</td>
</tr>
<tr>
<td>Radiation</td>
<td>Whole breast radiation may be added to lumpectomy.</td>
<td>Radiation to whole breast and lymph nodes if involved; follows chemotherapy if provided.</td>
<td>Radiation to chest wall and lymph nodes.</td>
<td>Selective radiation to bone or brain metastases.</td>
</tr>
<tr>
<td>Chemotherapy†</td>
<td>None</td>
<td>Systemic adjuvant therapy as indicated by ER, PR, and HER2 status and predictive tests for chemotherapy benefit.</td>
<td>• Pre-operative systemic therapy.</td>
<td>• If bone disease present, denosumab, zoledronic acid, or pamidronate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 1-year therapy with trastuzumab if HER2-positive.</td>
<td>• If ER and PR-negative; or ER and/or PR-positive and endocrine refractory; consider chemotherapy.‡</td>
</tr>
<tr>
<td>Endocrine</td>
<td>If ER-positive, consider tamoxifen for 5 years for prevention.</td>
<td>If ER-positive, tamoxifen for 10 years or aromatase inhibitor for 5 years (if post-menopausal only) or switching strategy of tamoxifen/aromatase inhibitor.</td>
<td>If ER-positive, tamoxifen for 10 years or aromatase inhibitor for 5 years (if post-menopausal only) or switching strategy of tamoxifen/aromatase inhibitor.</td>
<td>• Treatment regimen based on receptor status.</td>
</tr>
<tr>
<td>treatment§‡</td>
<td></td>
<td></td>
<td></td>
<td>• If ER positive, consider ovarian ablation/ suppression for premenopausal women</td>
</tr>
</tbody>
</table>

*Adapted from 2014 National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology.35
†Neoadjuvant/adjuvant chemotherapy: HER2-negative disease=AC (doxorubicin/cyclophosphamide) followed by paclitaxel, or TC (docetaxel and cyclophosphamide); HER2-positive disease=doxorubicin/cyclophosphamide followed by paclitaxel plus trastuzumab ± pertuzumab, or TCH (docetaxel/carboplatin/trastuzumab) ± pertuzumab.
‡Chemotherapy regimens for stage IV (metastatic cancer): preferred single agents=anthracyclines (doxorubicin, pegylated liposomal doxorubicin), taxanes (paclitaxel, docetaxel, nab-paclitaxel), anti-metabolites (capecitabine, gemcitabine), other microtubule inhibitors (vinorelbine, eribulin); chemotherapy combinations=CAF/FAC (cyclophosphamide/doxorubicin/fluorouracil), FEC (fluorouracil/epirubicin/cyclophosphamide), AC (doxorubicin/cyclophosphamide), EC (epirubicin/cyclophosphamide), CMF (cyclophosphamide/methotrexate/fluorouracil), docetaxel/capecitabine, GT (gemcitabine/paclitaxel), gemcitabine/carboplatin, paclitaxel/bevacizumab.
§Endocrine therapy for systemic disease: stage I-III=non-steroidal aromatase inhibitor (anastrozole, letrozole); steroidal aromatase inactivator ( exemestane); tamoxifen; stage IV=non-steroidal aromatase inhibitor (anastrozole, letrozole); steroidal aromatase inactivator including exemestane, exemestane + everolimus, fulvestrant, tamoxifen or toremifene, megestrol acetate, fluoxymesterone, ethinyl estradiol.

**Abbreviations:** DCIS=ductal carcinoma in situ; ER=estrogen receptor; HER2=human epidermal growth factor receptor 2; PR=progesterone receptor.
### Table 3. Imaging Modalities for Breast Cancer Screening of Average-Risk Women

<table>
<thead>
<tr>
<th>Imaging Modality†</th>
<th>Description; indication for use; average radiation dose (MQSA)‡</th>
<th>Limitations</th>
<th>Summary of performance‡</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mammography</strong></td>
<td>• A screening mammogram is performed in a woman with no clinical symptoms or complaints to detect early stage or clinically occult breast cancer. • Two views (craniocaudal and mediolateral oblique) of each breast are obtained for routine evaluation.</td>
<td>Limitations vary by type of mammography</td>
<td>Variable performance by type of mammography</td>
</tr>
<tr>
<td>Film mammography</td>
<td>• Uses x-rays transmitted through the breast tissue to create an image that is processed and displayed as a grayscale image directly on a film. • Adequate breast compression is required. • Women with larger breasts may require more than two views of each breast to ensure imaging of all breast tissues. • Average radiation dose is 4.7 mGy.</td>
<td>• Limited sensitivity in women with radiographically dense breasts. • Subject to artifacts from processing and storage. • Inability to manipulate the image following exposure.</td>
<td>All women: • Sensitivity 0.41 ±0.03 • Specificity 0.98 ±0.001 • PPV 0.13 ±0.01 Women &lt;50: • Sensitivity 0.35 ±0.06 • Specificity 0.98 ±0.001 • PPV 0.07 ±0.01</td>
</tr>
<tr>
<td>Digital mammography (DM)</td>
<td>• Digital detectors convert the x-ray photons to an electronic signal that is changed to a digital image and is processed and displayed as a grayscale image to be stored or sent electronically. Software can be used to help interpret digital images. • Available in &gt;90% of imaging centers in the United States as of 2013.²³⁸ • May be more effective than film in women &lt;50; woman with heterogeneously or extremely dense breast tissue; or pre- or perimenopausal women. • Average radiation dose is 3.7 mGy.</td>
<td>• Less spatial resolution compared with film. • More expensive (1.5 to 4 times cost of film).²³⁹</td>
<td>All women: • Sensitivity 0.41 ±0.03 • Specificity 0.98 ±0.001 • PPV 0.12 ±0.01 Women &lt;50: • Sensitivity 0.49 ±0.06 • Specificity 0.97 ±0.001 • PPV 0.08 ±0.01</td>
</tr>
<tr>
<td>Tomosynthesis</td>
<td>• A modification of DM that acquires images of a stationary, compressed breast at multiple angles during a short scan. • Individual images are reconstructed to generate a series of thin sections of images that can be displayed individually or in a loop. • Used in combination with standard DM for screening. Average radiation dose 1 to 2 times DM.</td>
<td>When performed in the screening setting, the patient is exposed to approximately twice the usual radiation dose, which can be even greater if the patient has dense or thick breasts.</td>
<td>Compared to digital mammography: PPV for recall 6.4% (±2.1%; 95% CI,1.7% to 2.5%, p&lt;0.001)²¹²</td>
</tr>
<tr>
<td>Imaging modality†</td>
<td>Description; indication for use; average radiation dose (MQSA)‡</td>
<td>Limitations</td>
<td>Summary of performance‡</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>• Sound waves used to create images of the breast using a non-invasive, hand held device. Images obtained by radiologist or technologist and are operator dependent. Whole breast ultrasound recently approved by the FDA for screening of patients with dense breasts. • Not currently indicated for routine screening. • There are no RCTs showing survival benefit of screening women with dense breasts with supplemental whole breast ultrasound screening (whole breast ultrasound) in addition to mammography. • Several states have now implemented standard reporting on breast density, which includes the recommendation for ultrasound for dense breasts. • No radiation.</td>
<td>• Not an appropriate initial screening modality for breast cancer, but has been approved as an adjunct to mammography for screening in women with increased breast density.</td>
<td>No data for average risk women; available performance measures are based on studies of women with increased risk and dense breasts.</td>
</tr>
<tr>
<td>MRI with and without contrast</td>
<td>• Magnetic fields are used to create an image of the breast. Intravenous contrast agent given for the procedure. • Not indicated for screening in average-risk populations; diagnostic modality in specific subpopulations. • No radiation.</td>
<td>Not an appropriate initial screening modality for breast cancer, but has been promoted as a screening test among women at elevated risk, including BRCA1/2 mutation carrier, strong family history of breast cancer, or several genetic syndromes.</td>
<td>No data for average risk women; available performance measures are based on studies of high-risk women.</td>
</tr>
</tbody>
</table>

*Average-risk women: women with <15% lifetime risk of breast cancer. Performance measures may vary based on risk, including breast density.
†Does not include other technologies not approved for screening: Positron emission mammography (PEM) and breast specific gamma imaging (BSGI).

Abbreviations: BRCA=breast cancer susceptibility gene; CI=confidence interval; DCIS=ductal carcinoma in situ; DM=digital mammography; FDA=U.S. Food and Drug Administration; mGy=milligray; MQSA=Mammography Quality Standards Act; MRI=magnetic resonance imaging; PPV=positive predictive value; RCT=randomized controlled trial.
### Table 4. Mammography Screening Trials

<table>
<thead>
<tr>
<th>Trial (references)</th>
<th>Year trial began</th>
<th>Setting/population (screening, n; control, n)*</th>
<th>Method of randomization</th>
<th>Comparison groups</th>
<th>Interval, months</th>
<th>Rounds, n</th>
<th>Views, n</th>
<th>Adherence, %</th>
<th>Duration, years</th>
<th>Longest followup, years</th>
<th>USPSTF quality rating; limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIP90,95,128, 131</td>
<td>1963</td>
<td>New York health plan members age 40-64 (30,239; 30,765)</td>
<td>Age and family size stratified pairs of women were individually randomized by drawing from a list</td>
<td>M + CBE vs. UC</td>
<td>12</td>
<td>4</td>
<td>2</td>
<td>46</td>
<td>4</td>
<td>18</td>
<td>Fair†‡§</td>
</tr>
<tr>
<td>CNBSS-1, CNBSS-2 [69, 76, 78]</td>
<td>1980</td>
<td>Self-selected participants from 15 centers in Canada age 40-49 (CNBSS-1; 25,214; 25,216) and 50-59 (CNBSS-2; 19,711; 19,694)</td>
<td>Individual within blocks stratified by center and 5-year age group after CBE</td>
<td>M + 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>
<td>4.5</td>
<td>25</td>
<td>Fair†</td>
</tr>
<tr>
<td>Gothenburg 82,241,242</td>
<td>1982</td>
<td>All women age 39-59 born between 1923 to 1944 living in Gothenburg, Sweden (21,650; 29,961)</td>
<td>Cluster, based on day of birth for 1923 to 1935 cohort (18%), by individual for 1936 to 1944 cohort (82%)</td>
<td>M vs. UC; controls offered screening after 5 years, trial completed after approximately 7 years</td>
<td>18</td>
<td>5</td>
<td>1-2</td>
<td>75</td>
<td>9</td>
<td>12</td>
<td>Fair†‡¶</td>
</tr>
<tr>
<td>Stockholm [243] 1981</td>
<td>Residents age 40-64 from southeast greater Stockholm, Sweden (40,318; 19,943)</td>
<td>Individual, by day of month; ratio of screening to control group 2:1</td>
<td>M vs. UC; controls screened after 5 years</td>
<td>24-28</td>
<td>2</td>
<td>1</td>
<td>81</td>
<td>4.8</td>
<td>11.4</td>
<td>11.4</td>
<td>Fair†</td>
</tr>
<tr>
<td>Malmö I &amp; II 81, 244,245 1976-1978</td>
<td>All women age 43-69 born between 1908 to 1945 living in Malmo, Sweden (MMST I=21,088; 21,195, MMST II=9,581; 8,212)</td>
<td>Individual, within birth year</td>
<td>M vs. UC; controls offered screening after year 14</td>
<td>18-24</td>
<td>9</td>
<td>1-2</td>
<td>70</td>
<td>10+</td>
<td>11-13; 15.5</td>
<td>Fair†</td>
<td></td>
</tr>
<tr>
<td>Swedish Two-County 83,246, 247 1977</td>
<td>Women age 40-70 from Ostergotland and Kopparberg counties in Sweden (77,080; 55,985)</td>
<td>Clusters, based on geographic units; blocks designed to be demographically homogeneous</td>
<td>M vs. UC; controls offered screening after year 7</td>
<td>24-33</td>
<td>3</td>
<td>1</td>
<td>84</td>
<td>7</td>
<td>20; 15.5</td>
<td>Fair†</td>
<td></td>
</tr>
</tbody>
</table>
### Table 4. Mammography Screening Trials

<table>
<thead>
<tr>
<th>Trial (references)</th>
<th>Year trial began</th>
<th>Setting/population (screening, n; control, n)*</th>
<th>Method of randomization</th>
<th>Comparison groups</th>
<th>Interval, months</th>
<th>Rounds, n</th>
<th>Views, n</th>
<th>Adherence, %</th>
<th>Duration, years</th>
<th>Longest followup, years</th>
<th>USPSTF quality rating; limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age‡¶§║1991</td>
<td>1991</td>
<td>Women age 39-41 from 23 National Health Service breast screening units in England, Scotland, and Wales (53,884; 106,956)</td>
<td>Individual stratified by general practitioner group with random number generation 1991 to 1992; 1992 onwards randomization via Health Authority computer system</td>
<td>M vs. UC; all women offered screening at ages 50-52</td>
<td>12</td>
<td>4-6, varied by center</td>
<td>2</td>
<td>57</td>
<td>9</td>
<td>13</td>
<td>Fair†‡ ¶‡§║</td>
</tr>
</tbody>
</table>

*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 followup.
‡ Important differential loss to followup or overall high loss to followup; adherence <80%.
§ Numbers of participants in screening and control groups vary by publication.
║ New data since prior recommendation.
¶ Maintenance of comparable groups (includes attrition, crossovers, adherence, contamination).

**Abbreviations:** BSE=breast self-examination; CBE=clinical breast examination; CNBSS=Canadian National Breast Screening Studies; HIP=Health Insurance Plan of New York; M=mammography; MMST=Malmö Mammographic Screening Trial; n=number; UC=usual care; USPSTF=U.S. Preventive Services Task Force; vs.=versus.
Table 5. Age-Specific Rates of Breast Cancer Mortality Reduction With Screening

Number of deaths prevented if 10,000 women were followed for 10 years

<table>
<thead>
<tr>
<th>Age, years</th>
<th>Mortality rate in the control group per 100,000 person-years (95% CI)*</th>
<th>Breast cancer mortality reduction RR (95% CI)†</th>
<th>Deaths prevented with screening over 10 years (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long case accrual</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39-49</td>
<td>34 (26 to 44)</td>
<td>0.88 (0.73 to 1.003)</td>
<td>4.1 (-0.1 to 9.3)</td>
</tr>
<tr>
<td>50-59</td>
<td>54 (50 to 58)</td>
<td>0.86 (0.68 to 0.97)</td>
<td>7.7 (1.6 to 17.2)</td>
</tr>
<tr>
<td>60-69</td>
<td>65 (52 to 81)</td>
<td>0.67 (0.54 to 0.83)</td>
<td>21.3 (10.7 to 31.7)</td>
</tr>
<tr>
<td>70-74</td>
<td>62 (48 to 80)</td>
<td>0.80 (0.51 to 1.28)</td>
<td>12.5 (-17.2 to 32.1)</td>
</tr>
<tr>
<td>50-69</td>
<td>58 (55 to 62)</td>
<td>0.78 (0.68 to 0.90)</td>
<td>12.5 (5.9 to 19.5)</td>
</tr>
<tr>
<td>Short case accrual</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39-49</td>
<td>23 (16 to 32)</td>
<td>0.84 (0.70 to 1.002)</td>
<td>3.5 (-0.1 to 7.4)</td>
</tr>
<tr>
<td>50-59</td>
<td>31 (24 to 39)</td>
<td>0.86 (0.69 to 1.007)</td>
<td>4.5 (-0.2 to 9.8)</td>
</tr>
<tr>
<td>60-69</td>
<td>40 (28 to 56)</td>
<td>0.67 (0.55 to 0.91)</td>
<td>12.1 (3.4 to 20.7)</td>
</tr>
<tr>
<td>70-74</td>
<td>49 (36 to 64)</td>
<td>0.90 (0.46 to 1.78)</td>
<td>12.2 (-37.7 to 26.9)</td>
</tr>
<tr>
<td>50-69</td>
<td>32 (24 to 41)</td>
<td>0.81 (0.69 to 0.95)</td>
<td>6.1 (1.2 to 10.9)</td>
</tr>
</tbody>
</table>

*Based on trials of screening included in the meta-analysis.
†From meta-analysis of screening trials using two different methods of case accrual.

Abbreviations: CI=confidence interval; RR=relative risk.
<table>
<thead>
<tr>
<th>Study design</th>
<th>Description</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-trend</td>
<td>Compares changes in breast cancer mortality among populations in relation to the introduction of screening (before/after or ecologic).</td>
<td>• Applicability to current populations and settings may be low.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Mortality rates may be affected by changes in diagnosis and treatment over time.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Analysis assumes constancy over time.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• High risk for lead-time and length-time biases depending on the choice of comparison time periods.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Comparison groups based on age or location are not stable over time.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Opportunistic screening in the control group may dilute mortality estimates or screening effects.</td>
</tr>
<tr>
<td>Incidence-based mortality</td>
<td>Compares mortality rates of women screened or invited to screen with women not screened or invited. To reflect the incidence of breast cancer, rather than prevalence, these studies include only breast cancer cases diagnosed during a specific time period that follows the initial screen.</td>
<td>• High risk for lead-time and length-time biases.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Short case accrual or followup periods inadequately determine mortality effect.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Opportunistic screening in the control group may dilute mortality estimates or screening effects.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Self-selection bias results in important differences between women attending screening and those who do not; including social, demographic, and health factors that independently influence outcomes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Dependent on correct choices of comparison groups.</td>
</tr>
<tr>
<td>Case-control</td>
<td>Compares histories of screening between women dying of breast cancer with women not dying of breast cancer.</td>
<td>• Self-selection bias.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Women who had access to screening likely had access to effective treatment.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Retrospective data analysis is subject to recall bias and missing data.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Lower power to detect mortality differences between groups.</td>
</tr>
<tr>
<td>Author, year</td>
<td>Study design</td>
<td>Population; age, year; participants, n</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>Coldman et al, 2008&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Time-trend</td>
<td>British Columbia, Canada, 4 cohorts based on date and age at first screening; 40-79; 658,151</td>
</tr>
<tr>
<td>Coldman et al, 2014&lt;sup&gt;24&lt;/sup&gt;</td>
<td>Incidence-based mortality</td>
<td>Canadian Screening Programs; 40-79; 2,796,472</td>
</tr>
<tr>
<td>Hellquist et al, 2011&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Prospective cohort (Poisson distribution)</td>
<td>Swedish counties in Mammography Screening of Young Women cohort; 40-49; 620,620</td>
</tr>
<tr>
<td>Hofvind et al, 2012&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Time-trend</td>
<td>Norwegian Breast Cancer Screening Program; 55-74; N=10,478 cancer cases</td>
</tr>
<tr>
<td>Hofvind et al, 2013&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Prospective cohort</td>
<td>Norwegian Breast Cancer Screening Program; 50-69; 699,628</td>
</tr>
<tr>
<td>Author, year</td>
<td>Study design</td>
<td>Population; age, year; participants, n</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>Mook et al, 2011&lt;sup&gt;127&lt;/sup&gt;</td>
<td>Retrospective cohort</td>
<td>Netherlands; 50-69; 2,592</td>
</tr>
<tr>
<td> </td>
<td> </td>
<td> </td>
</tr>
<tr>
<td> </td>
<td> </td>
<td> </td>
</tr>
<tr>
<td> </td>
<td> </td>
<td> </td>
</tr>
<tr>
<td> </td>
<td> </td>
<td> </td>
</tr>
<tr>
<td>Parvinen et al, 2011&lt;sup&gt;130&lt;/sup&gt;</td>
<td>Retrospective cohort</td>
<td>Finland, national screening program registry data; 40-49; 14,765</td>
</tr>
<tr>
<td> </td>
<td> </td>
<td> </td>
</tr>
<tr>
<td> </td>
<td> </td>
<td> </td>
</tr>
<tr>
<td>Schonberg et al, 2009&lt;sup&gt;34&lt;/sup&gt;</td>
<td>Retrospective cohort</td>
<td>U.S., medical record review at community health centers; &gt;80; 2,011</td>
</tr>
</tbody>
</table>

*Did not maintain comparable groups (includes attrition, crossovers, adherence, contamination).
†Statistical limitations including low power to detect differences.

**Abbreviations:** CI=confidence interval; HR=hazard ratio; MR=mortality ratio; n=number; NA=not applicable (quality rating criteria not available for this study design); NNS=number needed to screen; NR=not reported; RR=relative risk; U.S.=United States; vs.=versus.
Table 8. Advanced Breast Cancer Outcomes Reported in Screening Trials

<table>
<thead>
<tr>
<th>Trial (reference)</th>
<th>Stage</th>
<th>+Lymph nodes, n*</th>
<th>Size, mm†</th>
<th>Definition of advanced cancer‡</th>
<th>RR for advanced cancer (95% CI)§</th>
<th>Definition of advanced cancer║</th>
<th>RR for advanced cancer (95% CI)§</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIP131,132</td>
<td>I, II, III, IV</td>
<td>NR</td>
<td>NR</td>
<td>Stage II+</td>
<td>40-64 years: 0.85 (0.69 to 1.05)*</td>
<td>Stage III-IV</td>
<td>40-49 years: 0.87 (0.48 to 1.58)</td>
</tr>
<tr>
<td>CNBSS-175,76</td>
<td>NR</td>
<td>0, 1-3, 4+</td>
<td>1-9, 10-14, 15-19, 20-39, ≥40</td>
<td>Size ≥20 mm 1+ lymph node</td>
<td>40-49 years: 0.97 (0.74 to 1.25)*</td>
<td>Size ≥40 mm 4+ lymph nodes</td>
<td>40-49 years: 1.18 (0.67 to 2.03)</td>
</tr>
<tr>
<td>CNBSS-277,78</td>
<td>NR</td>
<td>0, 1-3, 4+</td>
<td>1-9, 10-14, 15-19, 20-39, ≥40</td>
<td>Size ≥20 mm 1+ lymph node</td>
<td>50-59 years: 0.84 (0.65 to 1.07)*</td>
<td>Size ≥40 mm 4+ lymph nodes</td>
<td>50-59 years: 0.75 (0.38 to 1.46)</td>
</tr>
<tr>
<td>Gothenburg82</td>
<td>NR</td>
<td>0, 1+</td>
<td>NR</td>
<td>1+ lymph node</td>
<td>39-59 years: 0.80 (0.61 to 1.05)*</td>
<td>NR</td>
<td>50-59 years: 0.91 (0.55 to 1.49)</td>
</tr>
<tr>
<td>Stockholm80</td>
<td>0, I, II, III-IV</td>
<td>NR</td>
<td>NR</td>
<td>Stage II+</td>
<td>40-64 years: 0.88 (0.68 to 1.12)‡</td>
<td>Stage III+</td>
<td>40-64 years: 1.15 (0.59 to 2.07)</td>
</tr>
<tr>
<td>Malmö81</td>
<td>0, I, II, III-IV, II-IV</td>
<td>NR</td>
<td>NR</td>
<td>Stage II+</td>
<td>45-70 years: 0.83 (0.68 to 1.00)‡</td>
<td>Stage III+</td>
<td>45-70 years: 0.82 (0.56 to 1.20)</td>
</tr>
<tr>
<td>Swedish Two-County83,132</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>Stage II+</td>
<td>40-74 years: 0.69 (0.61 to 0.78)†</td>
<td>NR</td>
<td>40-49 years: 1.57 (0.63 to 3.94)</td>
</tr>
<tr>
<td>Age246</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 ≥20 mm 1+ lymph node</td>
<td>39-49 years: 0.88 (0.73 to 1.05)*</td>
<td>Size ≥50 mm 4+ lymph nodes</td>
<td>39-49 years: 0.85 (0.57 to 1.23)</td>
</tr>
</tbody>
</table>

*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.
‡Autier, 200965.
§Screening vs. control.
║Represents the highest category of disease reported by the trials.

**Abbreviations:** CI=confidence interval; CNBSS=Canadian National Breast Screening Studies; HIP=Health Insurance Plan of Greater New York; mm=millimeter; n=number; NR=not reported; RR=relative risk; vs.=versus.
Table 9. Observational Studies of Advanced Cancer Outcomes With Mammography Screening

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Population; age, years; participants, n</th>
<th>Study years; comparison</th>
<th>Outcome measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Screening Frequency Trial Group, 2002</td>
<td>RCT</td>
<td>U.K., 5 screening units in NHS Breast Screening Programme; 50-62 years; 76,022</td>
<td>1989 to 1996; annual screening vs. no screening during study period</td>
<td>Size &gt;20 mm; ≥1 positive node</td>
<td>Invasive: 235 vs. 208 Tumor size &gt;20 mm: 27% (63/233) vs. 34% (69/203), p&lt;0.05 ≥1 node positive: 34% (63/185) vs. 37% (61/166), p=0.50</td>
</tr>
<tr>
<td>Buseman et al, 2003</td>
<td>Case series</td>
<td>U.S., Kaiser Permanente; 42-49 years; 247</td>
<td>1994 to 2000; screened vs. unscreened</td>
<td>Stage II-IV; III-IV</td>
<td>• 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>
</tr>
<tr>
<td>Dittus et al, 2013</td>
<td>Case series</td>
<td>U.S., BCSC data, multisite; 40-74 years; 4,432</td>
<td>1996 to 2008; 1-year vs. 2-year screening intervals</td>
<td>Stage; size &gt;20 mm; node positive</td>
<td>OR (95% CI) for 2-year vs. 1-year interval: No statistically significant differences for stage, size, lymph node positive by weight status.</td>
</tr>
<tr>
<td>Fernández et al, 2013</td>
<td>Case series</td>
<td>Spain, breast cancer program and regular public health system; 50-69 years; 904</td>
<td>2002 to 2012; screened vs. non-screened</td>
<td>Node positive; ≥3 nodes positive; size &gt;20 mm</td>
<td>• Cancer detection rate: 3.8/1,000 (475/123,445) vs. 9.4/1,000 (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=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. 48.5% (181/373), p=0.001</td>
</tr>
<tr>
<td>Goel et al, 2007</td>
<td>Case series</td>
<td>U.S., Vermont Breast Cancer Surveillance System; &gt;40 years; 1,944</td>
<td>1994 to 2002; 1-year vs. 2-year screening intervals</td>
<td>Advanced either Stage IIIB+; size &gt;20 mm; ≥1 positive node</td>
<td>Advanced: 21% vs. 24%, p=0.262 No statistically significant differences by age</td>
</tr>
<tr>
<td>Hubbard et al, 2011</td>
<td>Case series</td>
<td>U.S., BCSC data, multisite; 40-59 years; 4,492</td>
<td>1996 to 2006; 1-year vs. 2-year screening intervals</td>
<td>Stage IIIB+</td>
<td>Adjusted proportion (95% CI) of cancer stage for 2-year vs 1-year intervals: • Stage III or IV for 40-49 years: 4.8 (1.3 to 8.4) • No statistically significant differences for other stages</td>
</tr>
<tr>
<td>Jensen et al, 2003</td>
<td>Case series</td>
<td>Denmark and Sweden; 50-69 years; 2,104</td>
<td>1996 to 1997; regions with mammography screening vs. regions without</td>
<td>Stage III-IV; median size</td>
<td>• Stage III or IV: 8.8% (81/917) vs. 13.6% (162/1,187), p&lt;0.001 • Median tumor size (mm): 18 (Malmo) and 17 (Funen) vs. 20 (Aarhus and Northern Jutland), p&lt;0.001</td>
</tr>
<tr>
<td>Kerlikowske et al, 2013</td>
<td>Case series</td>
<td>U.S., BCSC data, multisite; 40-74 years; 11,474</td>
<td>1996 to 2008; 1-year vs. 2-year vs. 3-year screening intervals</td>
<td>Stage IIIB-IV</td>
<td>Adjusted OR (95% CI) for 2-year vs. 1-year intervals: • Stages IIIB-IV in age 40-49 + extreme breast density: 1.89 (1.06 to 3.39) • Tumor size &gt;20 mm in age 40-49 + extreme breast density: 2.39 (1.37 to 4.18) • No statistically significant differences for 50-74 years, 40-49 years without extreme density, or for any comparisons between 3-year vs. 2-year intervals</td>
</tr>
<tr>
<td>Author, year</td>
<td>Study design</td>
<td>Population; age, years; participants, n</td>
<td>Study years; comparison</td>
<td>Outcome measures</td>
<td>Results</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>----------------------------------------</td>
<td>------------------------</td>
<td>-----------------</td>
<td>---------</td>
</tr>
</tbody>
</table>
| Olivotto et al., 1999<sup>38</sup> | Case series | Canada, Screening Mammography Program of British Columbia; 40-89 years; 13,636 | 1989 to 1996; screening attenders vs. non-attenders | Stage III-IV; size >20 mm | • Invasive: 88% (1,712/1,946) vs. 92.3% (7,523/8,149), p<0.001  
• Stage III or IV: 4.3% (84/1,946) vs. 11.9% (969/8,149), p<0.001  
• Tumor size >20 mm: 24.1% (3413/1,946) vs. 38.3% (2,885/8,149), p<0.001 |
| Olsson et al., 2009<sup>39</sup> | Case series | Sweden, MMST; 45-69 years; 2478 | 1961 to 1991; invited to screen vs. not invited | Size >20 mm; node positive | • Tumor size >20 mm: 23% vs. 36%, p<0.05  
• Lymph node positive: 28% vs. 36%, p<0.05 |
| White et al, 2004<sup>45</sup> | Case series | U.S., BCSC data, multisite; 40-89 years; 7,840 | 1996 to 2001; 1-year vs. 2-year screening intervals | Stage III-IV; size >20 mm | • Stage III or IV: 3% vs. 4%  
• Tumor size >20 mm: 22% vs. 24%  
OR (95% CI) for 2-year interval vs. 1-year interval:  
• Late stage for invasive cancers only: 0.97 (0.84 to 1.13)  
• Tumor size >20 mm for invasive: 1.07 (0.92 to 1.24) |

**Abbreviations:** BCSC=Breast Cancer Surveillance Consortium; BMI=body mass index; CI=confidence interval; DCIS=ductal carcinoma in situ; kg=kilogram; m=meter; MMST=Malmö Mammographic Screening Trial; N=number; NHS=National Health Service; NR=not reported; NS=not statistically significant; U.S.=United States; U.K.=United Kingdom; vs.=versus.
<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Population; age, years; participants, n</th>
<th>Study years; comparison</th>
<th>Results</th>
</tr>
</thead>
</table>
| Buseman et al, 2003<sup>35</sup> | Case series | U.S., Kaiser Permanente; 42-49 years; 247 | 1994 to 2000; screened vs. nonscreened | Lumpectomy + radiation treatment: 61% (64/105) vs. 57% (81/142), NS  
Chemotherapy: 55% (58/105) vs. 61% (86/142), NS |
| Fernández et al, 2013<sup>36</sup> | Case series | Spain, breast cancer program and regular public health system; 50-69 years; 904 | 2002 to 2012; screened vs. nonscreened | Primary treatment:  
Conservative surgery: 83% (433/523) vs. 57% (230/403), p<0.001  
Radical surgery: 16% (84/523) vs. 41% (163/403), p<0.001  
Chemotherapy: 0.4% (2/510) vs. 0.8% (3/394), p<0.001  
Sentinel node biopsy: 73% (384/523) vs. 50% (200/403), p<0.001  
Adjuvant treatment:  
Chemotherapy: 41% (211/510) vs. 72% (284/394), p<0.001  
Hormone therapy: 86% (439/510) vs. 80% (317/394), p<0.001  
Radiotherapy: 87% (444/510) vs. 75% (296/393), p<0.001 |
| Jensen et al, 2003<sup>37</sup> | Case series | Denmark and Sweden; 50-69 years; 2,104 | 1996 to 1997; regions with mammography screening vs. regions without | Mastectomy: 61% (556/917) vs. 85% (893/1,051), p<0.001  
Lumpectomy: 32% (295/917) vs. 6.8% (72/1,051), p<0.001  
Biopsy only: 6.4% (59/917) vs. 8% (84/1,051), p<0.001 |
| Olivotto et al, 1999<sup>38</sup> | Case series | Canada, Screening Mammography Program of British Columbia; 40-89 years; 13,636 | 1989 to 1996; attenders vs. non-attenders | Definitive breast surgery:  
Total mastectomy: 35% (603/1,712) vs. 46% (3,452/7,523), p<0.001  
Breast conservation: 65% (1,109/1,712) vs. 54% (4,071/7,523), p<0.001  
Adjuvant systemic therapy:  
Tamoxifen alone: 29% (493/1,712) vs. 36% (2,694/7,523), p<0.001  
Chemotherapy: 23% (392/1,712) vs. 27% (2,060/7,523), p<0.001 |

**Abbreviations:** CI=confidence interval; MMST=Malmö Mammographic Screening Trial; N=number; NR=not reported; NS=not statistically significant; OR=odds ratio; U.S.=United States; U.K.=United Kingdom; vs.=versus.
### Table 11. Studies of Advanced Cancer Outcomes With Mammography Plus Tomosynthesis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Population; age, years; participants, n</th>
<th>Study years; comparison</th>
<th>Outcome measures/ definitions</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rose et al, 2013</td>
<td>Case series</td>
<td>U.S., multisite community-based breast center; &gt;18 years; 18,202 DM and 10,878 DM + T</td>
<td>2011 to 2012; DM vs. DM+T</td>
<td>Cancer detection rate; positive nodes</td>
<td>• Cancer detection rate: 4.0 vs 5.4/1000, NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Positive nodes: 4 vs. 6; p=0.84</td>
</tr>
<tr>
<td>Skaane et al, 2013</td>
<td>Post-intervention series</td>
<td>Norway, Oslo screening program; 50-69; 12,631</td>
<td>2010 to 2011; DM vs. DM+T (biennial screening)</td>
<td>Cancer detection rate; positive nodes; size &gt;20 mm</td>
<td>• Cancer detection rate: 6.1/1,000 vs. 8.0/1,000, (p=0.001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Positive nodes: 9 vs. 13, NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Size &gt;20 mm: 12 vs. 15, NS</td>
</tr>
</tbody>
</table>

**Abbreviations:** DM=digital mammography; mm= millimeter; NS=not statistically significant; T=tomosynthesis; U.S.=United States; vs.=versus.
Table 12. Age-Specific Rates of False-Positive and False-Negative Digital Mammography Results and Recommendations for Additional Imaging and Biopsies From a Single Screening Round in the BCSC

<table>
<thead>
<tr>
<th>Age, y</th>
<th>Women screened, n</th>
<th>Invasive breast cancer cases, n</th>
<th>DCIS cases, n</th>
<th>Outcomes, n per 1,000 women screened (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-49</td>
<td>113,770</td>
<td>349</td>
<td>191</td>
<td>False-positive mammography result: 121.2 (105.6, 138.7)</td>
</tr>
<tr>
<td>50-59</td>
<td>127,958</td>
<td>574</td>
<td>246</td>
<td>False-negative mammography result: 1.0 (0.9, 1.2)</td>
</tr>
<tr>
<td>60-69</td>
<td>94,507</td>
<td>651</td>
<td>208</td>
<td>Additional imaging recommended†: 124.9 (109.3, 142.3)</td>
</tr>
<tr>
<td>70-79</td>
<td>50,204</td>
<td>427</td>
<td>120</td>
<td>Biopsy recommended†: 16.4 (13.2, 20.3)</td>
</tr>
<tr>
<td>80-89</td>
<td>18,752</td>
<td>154</td>
<td>43</td>
<td>Screen-detected invasive cancer: 2.2 (1.8, 2.6)</td>
</tr>
</tbody>
</table>

*2-sided P-values and 95% confidence intervals from a logistic regression model that accounts for clustering by radiology facility using generalized estimating equations.
†After positive mammography result.

Abbreviations: CI=confidence interval; DCIS=ductal carcinoma in situ.
Table 13. Rates of False-Positive and False-Negative Digital Mammography Results and Recommendations for Additional Imaging and Biopsies Based on Time Since Last Mammography Examination*  

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time since last exam, mo</th>
<th>Age, y</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparing 9-18 vs. 19-30 months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women screened, n</td>
<td>9-18</td>
<td>79,637</td>
<td>91,864</td>
<td>71,324</td>
<td>39,474</td>
<td>14,865</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19-30</td>
<td>34,133</td>
<td>36,094</td>
<td>23,183</td>
<td>10,730</td>
<td>3,887</td>
<td></td>
</tr>
<tr>
<td>Invasive breast cancer cases, n</td>
<td>9-18</td>
<td>240</td>
<td>391</td>
<td>474</td>
<td>322</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19-30</td>
<td>109</td>
<td>183</td>
<td>177</td>
<td>105</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>DCIS cases, n</td>
<td>9-18</td>
<td>126</td>
<td>185</td>
<td>156</td>
<td>94</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19-30</td>
<td>65</td>
<td>61</td>
<td>52</td>
<td>26</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes, n per 1,000 women screened (95% CI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False-positive mammography result</td>
<td>9-18</td>
<td>122.1 (105.4, 141.0)</td>
<td>0.65</td>
<td>94.2 (83.3, 106.5)</td>
<td>0.37</td>
<td>80.6 (72.8, 89.2)</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>19-30</td>
<td>119.0 (103.0, 137.1)</td>
<td>90.5 (80.4, 101.8)</td>
<td>81.1 (71.4, 92.1)</td>
<td>71.6 (62.2, 82.2)</td>
<td>60.2 (49.3, 73.3)</td>
<td></td>
</tr>
<tr>
<td>False-negative mammography result</td>
<td>9-18</td>
<td>1.1 (0.9, 1.3)</td>
<td>0.14</td>
<td>1.2 (1.0, 1.4)</td>
<td>0.06</td>
<td>1.3 (1.0, 1.6)</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>19-30</td>
<td>0.8 (0.6, 1.1)</td>
<td>0.9 (0.6, 1.2)</td>
<td>0.9 (0.6, 1.5)</td>
<td>1.0 (0.5, 2.0)</td>
<td>0.8 (0.3, 2.3)</td>
<td></td>
</tr>
<tr>
<td>Additional imaging recommended†</td>
<td>9-18</td>
<td>125.6 (109.0, 144.3)</td>
<td>0.74</td>
<td>99.3 (88.2, 111.7)</td>
<td>0.47</td>
<td>88.2 (80.2, 96.9)</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>19-30</td>
<td>123.3 (107.0, 141.7)</td>
<td>96.4 (85.9, 108.0)</td>
<td>90.1 (80.1, 101.2)</td>
<td>82.8 (72.5, 94.3)</td>
<td>71.3 (59.8, 84.8)</td>
<td></td>
</tr>
<tr>
<td>Biopsy recommended†</td>
<td>9-18</td>
<td>15.6 (12.8, 19.0)</td>
<td>0.11</td>
<td>15.7 (12.7, 19.3)</td>
<td>0.50</td>
<td>15.9 (14.0, 18.2)</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>19-30</td>
<td>18.2 (13.7, 24.1)</td>
<td>16.4 (12.5, 21.4)</td>
<td>18.4 (14.7, 23.0)</td>
<td>18.5 (14.6, 23.5)</td>
<td>18.3 (13.9, 24.0)</td>
<td></td>
</tr>
<tr>
<td>Screen-detected invasive cancer</td>
<td>9-18</td>
<td>2.0 (1.6, 2.5)</td>
<td>0.12</td>
<td>3.2 (2.7, 3.7)</td>
<td>0.009</td>
<td>5.5 (4.9, 6.2)</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>19-30</td>
<td>2.5 (2.1, 3.0)</td>
<td>4.3 (3.7, 5.1)</td>
<td>6.8 (6.7, 8.1)</td>
<td>8.9 (7.4, 10.8)</td>
<td>8.2 (5.8, 11.6)</td>
<td></td>
</tr>
<tr>
<td>Screen-detected DCIS</td>
<td>9-18</td>
<td>1.5 (1.2, 1.8)</td>
<td>0.18</td>
<td>1.9 (1.5, 2.4)</td>
<td>0.13</td>
<td>2.0 (1.8, 2.4)</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>19-30</td>
<td>1.8 (1.4, 2.3)</td>
<td>1.6 (1.2, 2.0)</td>
<td>2.2 (1.4, 3.3)</td>
<td>2.2 (1.4, 3.6)</td>
<td>2.8 (1.5, 5.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Comparing 11-14 vs. 23-26 months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women screened, n</td>
<td>11-14</td>
<td>55,278</td>
<td>65,219</td>
<td>53,419</td>
<td>30,497</td>
<td>11,299</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23-26</td>
<td>13,584</td>
<td>14,407</td>
<td>9,907</td>
<td>4,291</td>
<td>1,504</td>
<td></td>
</tr>
<tr>
<td>Invasive breast cancer cases, n</td>
<td>11-14</td>
<td>163</td>
<td>274</td>
<td>348</td>
<td>247</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23-26</td>
<td>42</td>
<td>70</td>
<td>76</td>
<td>41</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>DCIS cases, n</td>
<td>11-14</td>
<td>83</td>
<td>127</td>
<td>111</td>
<td>71</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23-26</td>
<td>26</td>
<td>22</td>
<td>23</td>
<td>12</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes, n per 1,000 women screened (95% CI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False-positive mammography result</td>
<td>11-14</td>
<td>119.1 (103.5, 136.8)</td>
<td>0.69</td>
<td>93.3 (82.8, 105.0)</td>
<td>0.46</td>
<td>79.2 (72.2, 86.8)</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>23-26</td>
<td>115.8 (98.7, 135.4)</td>
<td>89.9 (78.8, 102.4)</td>
<td>79.6 (70.3, 90.2)</td>
<td>65.7 (56.7, 76.0)</td>
<td>61.2 (47.3, 78.7)</td>
<td></td>
</tr>
<tr>
<td>False-negative mammography result</td>
<td>11-14</td>
<td>1.2 (1.0, 1.5)</td>
<td>0.20</td>
<td>1.2 (1.0, 1.5)</td>
<td>0.11</td>
<td>1.2 (0.9, 1.6)</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>23-26</td>
<td>0.9 (0.5, 1.5)</td>
<td>0.8 (0.4, 1.4)</td>
<td>0.8 (0.4, 1.8)</td>
<td>1.4 (0.6, 3.4)</td>
<td>2.0 (0.7, 6.0)</td>
<td></td>
</tr>
</tbody>
</table>
Table 13. Rates of False-Positive and False-Negative Digital Mammography Results and Recommendations for Additional Imaging and Biopsies Based on Time Since Last Mammography Examination*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time since last exam, mo</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional imaging recommended†</td>
<td>11-14</td>
<td>122.4</td>
<td>98.3</td>
<td>86.6</td>
<td>76.6</td>
<td>71.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(106.7, 139.9)</td>
<td>(87.7, 109.9)</td>
<td>(79.5, 94.3)</td>
<td>(69.3, 84.5)</td>
<td>(65.4, 77.7)</td>
</tr>
<tr>
<td></td>
<td>23-26</td>
<td>119.9</td>
<td>95.5</td>
<td>88.8</td>
<td>76.7</td>
<td>71.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(102.6, 139.7)</td>
<td>(83.9, 108.5)</td>
<td>(79.2, 99.5)</td>
<td>(66.5, 88.2)</td>
<td>(57.1, 88.3)</td>
</tr>
<tr>
<td>Biopsy recommended†</td>
<td>11-14</td>
<td>14.7</td>
<td>15.1</td>
<td>15.2</td>
<td>16.6</td>
<td>13.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(12.2, 17.8)</td>
<td>(12.2, 18.6)</td>
<td>(13.5, 17.2)</td>
<td>(14.5, 18.9)</td>
<td>(10.8, 16.0)</td>
</tr>
<tr>
<td></td>
<td>23-26</td>
<td>16.9</td>
<td>15.8</td>
<td>18.8</td>
<td>17.0</td>
<td>16.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11.9, 24.0)</td>
<td>(11.7, 21.3)</td>
<td>(15.2, 23.2)</td>
<td>(12.6, 23.0)</td>
<td>(11.2, 24.7)</td>
</tr>
<tr>
<td>Screen-detected invasive cancer</td>
<td>11-14</td>
<td>1.8</td>
<td>3.1</td>
<td>5.5</td>
<td>6.8</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.5, 2.3)</td>
<td>(2.6, 3.7)</td>
<td>(4.9, 6.2)</td>
<td>(5.8, 7.9)</td>
<td>(4.5, 7.8)</td>
</tr>
<tr>
<td></td>
<td>23-26</td>
<td>2.3</td>
<td>4.2</td>
<td>7.0</td>
<td>8.4</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.6, 3.2)</td>
<td>(3.3, 5.4)</td>
<td>(5.7, 8.5)</td>
<td>(5.7, 12.4)</td>
<td>(4.9, 13.1)</td>
</tr>
<tr>
<td>Screen-detected DCIS</td>
<td>11-14</td>
<td>1.4</td>
<td>1.9</td>
<td>1.9</td>
<td>2.2</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.1, 1.8)</td>
<td>(1.4, 2.4)</td>
<td>(1.6, 2.3)</td>
<td>(1.6, 3.0)</td>
<td>(0.9, 2.8)</td>
</tr>
<tr>
<td></td>
<td>23-26</td>
<td>1.8</td>
<td>1.4</td>
<td>2.2</td>
<td>2.6</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.3, 2.7)</td>
<td>(0.9, 2.1)</td>
<td>(1.3, 3.7)</td>
<td>(1.3, 5.1)</td>
<td>(0.6, 6.1)</td>
</tr>
</tbody>
</table>

*2-sided P-values and 95% confidence intervals from a logistic regression model that accounts for clustering by radiology facility using generalized estimating equations.
†After positive mammography result.

**Abbreviations:** CI=confidence interval; DCIS=ductal carcinoma in situ.
Table 14. Rates of False-Positive Results After Screening With Digital Mammography by Risk Factors

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>40-49 Age, y</th>
<th>50-59 Age, y</th>
<th>60-69 Age, y</th>
<th>70-79 Age, y</th>
<th>80-89 Age, y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women screened, n</td>
<td>113,770</td>
<td>127,958</td>
<td>94,507</td>
<td>50,204</td>
<td>18,752</td>
</tr>
<tr>
<td>False-positive, n</td>
<td>13,784</td>
<td>11,923</td>
<td>7,633</td>
<td>3,494</td>
<td>1,223</td>
</tr>
<tr>
<td>Number per 1,000 women screened per round (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-degree relatives with breast cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>118.7 (104.3, 134.7)</td>
<td>90.4 (81.1, 100.7)</td>
<td>0.005 79.4 (71.8, 87.7)</td>
<td>0.02 68.6 (61.1, 76.8)</td>
<td>0.11 63.3 (56.8, 70.5)</td>
</tr>
<tr>
<td>One or more</td>
<td>139.8 (113.9, 170.5)</td>
<td>109.0 (92.3, 128.2)</td>
<td>0.005 87.2 (77.2, 98.4)</td>
<td>0.02 75.0 (67.8, 83.1)</td>
<td>0.11 73.1 (64.1, 83.3)</td>
</tr>
<tr>
<td>Breast density†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat-Scattered</td>
<td>108.4 (95.5, 122.7)</td>
<td>&lt;0.001 80.5 (71.1, 90.9)</td>
<td>&lt;0.001 74.1 (66.4, 82.6)</td>
<td>&lt;0.001 67.3 (60.4, 74.9)</td>
<td>0.003 60.3 (54.0, 67.4)</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>142.2 (120.2, 167.4)</td>
<td>115.8 (100.3, 133.2)</td>
<td>0.005 101.8 (91.0, 113.8)</td>
<td>0.005 88.7 (78.7, 99.9)</td>
<td>0.005 82.4 (72.6, 93.5)</td>
</tr>
<tr>
<td>Extreme</td>
<td>112.1 (94.4, 132.7)</td>
<td>92.7 (77.5, 110.5)</td>
<td>0.005 75.2 (64.7, 87.1)</td>
<td>0.005 57.7 (43.9, 75.5)</td>
<td>0.005 85.1 (61.7, 116.2)</td>
</tr>
<tr>
<td>Benign breast biopsy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>114.3 (99.8, 130.5)</td>
<td>0.001 85.9 (76.7, 96.0)</td>
<td>&lt;0.001 74.6 (66.8, 83.1)</td>
<td>&lt;0.001 63.4 (56.2, 71.3)</td>
<td>&lt;0.001 63.0 (56.3, 70.6)</td>
</tr>
<tr>
<td>Previous</td>
<td>167.3 (140.6, 197.9)</td>
<td>122.5 (106.2, 140.7)</td>
<td>0.006 98.6 (88.8, 109.3)</td>
<td>0.006 88.6 (79.1, 99.2)</td>
<td>0.006 71.6 (62.3, 82.3)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>127.0 (115.5, 139.4)</td>
<td>0.001 97.6 (89.5, 106.4)</td>
<td>0.01 83.8 (77.4, 90.7)</td>
<td>0.006 73.5 (67.7, 79.8)</td>
<td>&lt;0.001 68.9 (62.6, 75.7)</td>
</tr>
<tr>
<td>Black</td>
<td>92.6 (82.0, 104.5)</td>
<td>78.9 (65.2, 95.3)</td>
<td>0.01 64.5 (53.6, 77.3)</td>
<td>0.01 58.9 (51.7, 67.0)</td>
<td>0.01 52.4 (43.6, 63.0)</td>
</tr>
<tr>
<td>Asian</td>
<td>85.2 (72.2, 100.4)</td>
<td>67.6 (56.5, 80.7)</td>
<td>0.01 58.0 (47.9, 70.2)</td>
<td>0.01 43.6 (36.9, 51.6)</td>
<td>0.01 35.8 (29.6, 43.4)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>125.4 (106.8, 146.7)</td>
<td>80.9 (69.1, 94.6)</td>
<td>0.01 72.9 (60.3, 87.8)</td>
<td>0.01 60.7 (50.6, 72.8)</td>
<td>0.01 55.7 (31.3, 97.2)</td>
</tr>
<tr>
<td>Other</td>
<td>127.8 (105.8, 153.6)</td>
<td>102.3 (88.5, 117.8)</td>
<td>0.01 91.5 (76.2, 109.5)</td>
<td>0.01 72.6 (53.3, 98.2)</td>
<td>0.01 48.9 (29.3, 80.6)</td>
</tr>
<tr>
<td>Menopausal hormone therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>123.3 (107.4, 141.2)</td>
<td>0.69 91.8 (81.6, 103.2)</td>
<td>0.27 76.2 (69.2, 84.0)</td>
<td>0.22 67.6 (61.1, 74.8)</td>
<td>0.01 62.2 (55.5, 69.8)</td>
</tr>
<tr>
<td>Combination</td>
<td>122.0 (78.8, 184.1)</td>
<td>131.1 (99.5, 170.7)</td>
<td>0.27 122.5 (87.3, 169.2)</td>
<td>0.27 105.9 (81.8, 136.0)</td>
<td>0.27 94.0 (74.0, 118.8)</td>
</tr>
<tr>
<td>Estrogen</td>
<td>108.7 (84.4, 138.8)</td>
<td>101.3 (87.1, 117.6)</td>
<td>0.27 97.6 (77.3, 122.5)</td>
<td>0.27 114.0 (94.8, 136.5)</td>
<td>0.27 89.1 (68.5, 115.1)</td>
</tr>
<tr>
<td>Oral contraceptives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No current</td>
<td>122.9 (107.2, 140.6)</td>
<td>0.05 93.6 (83.1, 105.4)</td>
<td>0.63 NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Current use</td>
<td>106.2 (86.4, 130.0)</td>
<td>97.0 (81.3, 115.2)</td>
<td>0.63 NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 14. Rates of False-Positive Results After Screening With Digital Mammography by Risk Factors*

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body mass index, kg/m²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>129.0 (113.8, 145.9)</td>
<td>99.5 (89.3, 110.8)</td>
<td>85.8 (77.9, 94.4)</td>
<td>70.5 (62.0, 80.0)</td>
<td>73.9 (60.6, 89.8)</td>
</tr>
<tr>
<td>25 to &lt;30</td>
<td>124.8 (110.1, 141.2)</td>
<td>93.6 (85.0, 103.0)</td>
<td>78.6 (69.5, 88.9)</td>
<td>72.7 (64.8, 81.6)</td>
<td>62.2 (51.4, 75.1)</td>
</tr>
<tr>
<td>≥30</td>
<td>107.2 (96.0, 119.5)</td>
<td>86.1 (77.7, 95.2)</td>
<td>81.1 (74.1, 88.6)</td>
<td>74.2 (64.1, 85.7)</td>
<td>73.8 (59.1, 91.9)</td>
</tr>
</tbody>
</table>

*2-sided P-values and 95% confidence intervals from a logistic regression model that accounts for clustering by radiology facility using generalized estimating equations.†Categories include: almost entirely fat=fat; scattered fibroglandular densities=scattered; heterogeneously dense=heterogeneous; and extremely dense=extreme.**Abbreviations:** kg=kilogram; m=meter; NA=not applicable; peri=perimenopausal; pre=premenopausal; post=postmenopausal.
Table 15. Rates of False-Negative Results After Screening With Digital Mammography by Risk Factors*

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Age, y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>40-49</td>
</tr>
<tr>
<td><strong>Women screened, n</strong></td>
<td>113,770</td>
</tr>
<tr>
<td><strong>False-negative mammography result, n</strong></td>
<td>115</td>
</tr>
<tr>
<td><strong>Number per 1,000 women screened per round (95% CI)</strong></td>
<td></td>
</tr>
<tr>
<td>First-degree relatives with breast cancer</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>0.9 (0.8, 1.1)</td>
</tr>
<tr>
<td>One or more</td>
<td>1.8 (1.3, 2.5)</td>
</tr>
<tr>
<td>Breast density</td>
<td></td>
</tr>
<tr>
<td>Fat-Scattered</td>
<td>0.4 (0.3, 0.6)</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>1.3 (1.0, 1.7)</td>
</tr>
<tr>
<td>Extreme</td>
<td>1.7 (1.2, 2.5)</td>
</tr>
<tr>
<td>Benign breast biopsy†</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>0.9 (0.8, 1.1)</td>
</tr>
<tr>
<td>Previous</td>
<td>1.1 (0.7, 1.7)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1.2 (1.0, 1.4)</td>
</tr>
<tr>
<td>Black</td>
<td>0.7 (0.3, 1.4)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.8 (0.5, 1.3)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.5 (0.2, 1.6)</td>
</tr>
<tr>
<td>Other</td>
<td>1.1 (0.4, 3.2)</td>
</tr>
<tr>
<td>Menopausal status</td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>1.2 (1.0, 1.4)</td>
</tr>
<tr>
<td>Peri</td>
<td>0.8 (0.2, 2.5)</td>
</tr>
<tr>
<td>Post</td>
<td>0.7 (0.4, 1.3)</td>
</tr>
<tr>
<td>Menopausal hormone therapy</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1.0 (0.9, 1.2)</td>
</tr>
<tr>
<td>Combination</td>
<td>0‡</td>
</tr>
<tr>
<td>Estrogen only</td>
<td>1.5 (0.2, 10.1)</td>
</tr>
<tr>
<td>Oral contraceptives</td>
<td></td>
</tr>
<tr>
<td>No current</td>
<td>1.0 (0.8, 1.2)</td>
</tr>
<tr>
<td>Current use</td>
<td>1.1 (0.6, 2.1)</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>1.4 (1.2, 1.7)</td>
</tr>
<tr>
<td>25 to &lt;30</td>
<td>0.8 (0.6, 1.3)</td>
</tr>
<tr>
<td>≥30</td>
<td>0.7 (0.3, 1.4)</td>
</tr>
</tbody>
</table>

*2-sided P-value and 95% confidence intervals from logistic regression model that accounts for clustering by radiology facility using generalized estimating equations.
†Categories include: almost entirely fat=fatty; scattered fibroglandular densities=scattered; heterogeneously dense=heterogeneous; and extremely dense=extreme.
‡No false-negative outcomes. Category omitted from model used to obtain CI and P-value.

**Abbreviations:** kg=kilogram; m=meter; NA=not applicable; peri=perimenopausal; pre=prenopausal; post=postmenopausal.
## Table 16. Rates of Recommendations for Additional Imaging After Screening With Digital Mammography by Risk Factors*

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th><strong>40-49</strong></th>
<th><strong>50-59</strong></th>
<th><strong>60-69</strong></th>
<th><strong>70-79</strong></th>
<th><strong>80-89</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women screened, n</strong></td>
<td>113,770</td>
<td>127,958</td>
<td>94,507</td>
<td>50,204</td>
<td>18,752</td>
</tr>
<tr>
<td><strong>Additional imaging recommended, n</strong></td>
<td>14,209</td>
<td>12,604</td>
<td>8,380</td>
<td>3,968</td>
<td>1,396</td>
</tr>
<tr>
<td><strong>Number per 1,000 women screened per round (95% CI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>First-degree relatives with breast cancer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>122.1 (107.7, 138.1)</td>
<td>95.2 (85.8, 105.6)</td>
<td>86.7 (79.0, 95.1)</td>
<td>77.5 (69.9, 85.7)</td>
<td>71.7 (64.6, 79.5)</td>
</tr>
<tr>
<td>One or more</td>
<td>145.6 (119.6, 176.2)</td>
<td>117.1 (99.7, 137.0)</td>
<td>98.3 (87.9, 109.8)</td>
<td>86.9 (79.1, 95.4)</td>
<td>86.0 (75.5, 97.7)</td>
</tr>
<tr>
<td><strong>Breast density</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat-Scattered</td>
<td>110.8 (97.9, 125.2)</td>
<td>84.4 (74.8, 95.1)</td>
<td>81.0 (73.1, 89.6)</td>
<td>76.6 (68.5, 83.4)</td>
<td>68.8 (61.7, 76.9)</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>146.0 (123.9, 171.3)</td>
<td>121.6 (105.8, 139.3)</td>
<td>110.6 (99.7, 122.6)</td>
<td>99.0 (87.9, 111.4)</td>
<td>93.6 (82.4, 106.2)</td>
</tr>
<tr>
<td>Extreme</td>
<td>116.5 (98.4, 137.4)</td>
<td>98.4 (83.1, 116.2)</td>
<td>81.0 (70.3, 93.2)</td>
<td>63.3 (49.7, 80.1)</td>
<td>92.0 (66.6, 125.7)</td>
</tr>
<tr>
<td><strong>Benign breast biopsy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>117.8 (103.4, 134.0)</td>
<td>90.9 (81.7, 101.0)</td>
<td>81.9 (74.1, 90.6)</td>
<td>72.2 (65.1, 79.9)</td>
<td>72.1 (64.3, 80.7)</td>
</tr>
<tr>
<td>Previous</td>
<td>172.5 (145.9, 202.8)</td>
<td>129.3 (112.8, 147.8)</td>
<td>108.2 (98.2, 118.9)</td>
<td>100.5 (90.0, 112.1)</td>
<td>82.7 (72.9, 93.7)</td>
</tr>
<tr>
<td><strong>Race/ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>131.1 (119.4, 143.8)</td>
<td>103.2 (94.8, 112.3)</td>
<td>92.4 (85.7, 99.4)</td>
<td>83.3 (77.1, 90.1)</td>
<td>78.0 (71.1, 85.5)</td>
</tr>
<tr>
<td>Black</td>
<td>95.9 (85.0, 108.0)</td>
<td>82.6 (68.4, 99.4)</td>
<td>70.8 (59.3, 84.3)</td>
<td>66.3 (59.1, 74.4)</td>
<td>60.3 (49.1, 74.0)</td>
</tr>
<tr>
<td>Asian</td>
<td>89.1 (76.0, 104.2)</td>
<td>73.5 (62.1, 86.8)</td>
<td>64.6 (54.0, 77.0)</td>
<td>52.6 (44.9, 61.4)</td>
<td>40.5 (33.4, 48.9)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>127.8 (109.2, 149.0)</td>
<td>84.6 (71.9, 99.3)</td>
<td>76.9 (64.1, 92.0)</td>
<td>72.1 (61.6, 84.3)</td>
<td>62.3 (38.7, 98.9)</td>
</tr>
<tr>
<td>Other</td>
<td>131.6 (109.8, 157.1)</td>
<td>109.8 (97.1, 123.8)</td>
<td>98.8 (82.5, 117.8)</td>
<td>84.7 (64.0, 111.3)</td>
<td>65.2 (39.4, 106.2)</td>
</tr>
<tr>
<td><strong>Menopausal status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>135.4 (117.4, 155.6)</td>
<td>124.6 (113.6, 136.4)</td>
<td>&lt;0.001 NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Peri</td>
<td>109.0 (92.8, 127.7)</td>
<td>101.4 (78.7, 129.8)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Post</td>
<td>114.2 (103.1, 126.4)</td>
<td>92.7 (84.0, 102.1)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Menopausal hormone therapy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>127.0 (111.2, 144.8)</td>
<td>97.0 (86.7, 108.5)</td>
<td>83.8 (76.5, 91.7)</td>
<td>76.5 (69.8, 83.9)</td>
<td>71.5 (64.2, 79.6)</td>
</tr>
<tr>
<td>Combination</td>
<td>125.8 (103.6, 185.0)</td>
<td>137.4 (105.5, 177.1)</td>
<td>129.5 (96.3, 172.0)</td>
<td>120.7 (94.9, 152.4)</td>
<td>106.6 (79.4, 141.6)</td>
</tr>
<tr>
<td>Estrogen only</td>
<td>110.1 (85.6, 140.7)</td>
<td>103.4 (90.9, 121.8)</td>
<td>106.1 (86.0, 130.3)</td>
<td>125.1 (106.4, 146.6)</td>
<td>108.4 (82.6, 136.1)</td>
</tr>
</tbody>
</table>
Table 16. Rates of Recommendations for Additional Imaging After Screening With Digital Mammography by Risk Factors*

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Age, y</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral contraceptives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No current</td>
<td>126.6</td>
<td>99.0</td>
<td>0.85</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Current use</td>
<td>110.4</td>
<td>100.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>133.9</td>
<td>105.9</td>
<td>93.4</td>
<td>79.5</td>
<td>83.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(118.1, 151.3)</td>
<td>(95.6, 117.2)</td>
<td>(85.4, 102.1)</td>
<td>(70.5, 89.5)</td>
<td>(69.8, 99.5)</td>
<td></td>
</tr>
<tr>
<td>25 to &lt;30</td>
<td>129.2</td>
<td>99.3</td>
<td>88.7</td>
<td>84.1</td>
<td>69.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(114.7, 145.2)</td>
<td>(90.4, 108.9)</td>
<td>(79.1, 99.4)</td>
<td>(75.5, 93.6)</td>
<td>(58.5, 82.3)</td>
<td></td>
</tr>
<tr>
<td>≥30</td>
<td>110.7</td>
<td>93.1</td>
<td>89.2</td>
<td>89.3</td>
<td>88.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(99.4, 123.2)</td>
<td>(84.2, 102.8)</td>
<td>(82.1, 96.8)</td>
<td>(78.5, 101.5)</td>
<td>(71.5, 108.8)</td>
<td></td>
</tr>
</tbody>
</table>

*2-sided P-values and 95% confidence intervals from a logistic regression model that accounts for clustering by radiology facility using generalized estimating equations.
†Categories include: almost entirely fat=fat; scattered fibroglandular densities=scattered; heterogeneously dense=heterogeneous; and extremely dense=extreme.

Abbreviations: kg=kilogram; m=meter; NA=not applicable; peri=perimenopausal; pre=premenopausal; post=postmenopausal.
Table 17. Rates of Recommendations for Biopsy After Screening With Digital Mammography by Risk Factors*

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women screened, n</td>
<td>113,770</td>
<td>127,958</td>
<td>94,507</td>
<td>50,204</td>
<td>18,752</td>
</tr>
<tr>
<td>Biopsy recommended, n</td>
<td>1,863</td>
<td>2,030</td>
<td>1,562</td>
<td>880</td>
<td>293</td>
</tr>
<tr>
<td><strong>Number per 1,000 women screened per round (95% CI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**First-degree relatives with breast cancer**

| None | 15.7 (12.6, 19.4) | <0.001 | 15.8 (13.7, 18.3) | 0.002 | 17.0 (14.7, 19.6) | 0.09 | 15.2 (12.8, 18.0) | 0.24 |
| One or more | 21.1 (16.9, 26.3) | <0.001 | 15.6 (13.7, 17.7) | 0.008 | 16.2 (14.2, 18.4) | 0.007 | 14.2 (12.0, 16.8) | 0.07 |

**Breast density†**

| Fat-Scattered | 12.2 (9.9, 15.0) | <0.001 | 11.8 (9.6, 14.5) | <0.001 | 15.6 (13.7, 17.7) | <0.001 | 14.2 (12.0, 16.8) | 0.54 |
| Heterogeneous | 18.9 (15.8, 22.5) | 20.2 (17.3, 23.7) | 19.3 (16.9, 22.2) | 21.0 (18.0, 24.5) | 19.0 (15.5, 23.2) | 16.1 (8.0, 32.1) |       |
| Extreme       | 20.2 (16.8, 24.3) | 19.2 (14.3, 25.7) | 13.8 (10.5, 18.2) | 13.0 (7.2, 23.3) | 16.1 (8.0, 32.1) |       |       |       |

**Benign breast biopsy**

| None | 14.8 (11.8, 18.7) | <0.001 | 13.9 (11.1, 17.3) | <0.002 | 15.0 (12.7, 17.8) | <0.001 | 15.3 (13.1, 17.7) | <0.001 | 15.8 (13.4, 18.7) | 0.54 |
| Previous | 27.8 (22.8, 33.7) | 25.1 (20.1, 31.2) | 21.8 (19.1, 24.9) | 25.2 (21.4, 29.7) | 17.1 (13.7, 21.5) |       |       |       |

**Race/ethnicity**

| White        | 16.7 (13.7, 20.3) | 0.21 | 16.6 (13.6, 20.2) | 0.39 | 17.6 (15.6, 20.0) | 0.05 | 18.7 (16.6, 21.2) | 0.23 | 16.2 (13.5, 19.4) | 0.12 |
| Black        | 13.6 (10.4, 17.8) | 14.7 (10.4, 20.6) | 13.9 (10.6, 18.0) | 14.9 (11.4, 19.5) | 8.9 (4.4, 18.0) |       |       |       |
| Asian        | 16.2 (10.6, 24.5) | 14.8 (9.5, 22.9) | 12.0 (6.9, 20.6) | 11.8 (6.8, 20.3) | 9.2 (5.6, 15.3) |       |       |       |
| Hispanic     | 16.3 (10.3, 25.6) | 11.9 (8.1, 17.5) | 14.2 (11.4, 17.6) | 15.9 (10.1, 25.1) | 16.4 (8.5, 31.5) |       |       |       |
| Other        | 19.8 (14.4, 27.3) | 17.4 (10.5, 28.6) | 16.4 (10.8, 24.8) | 16.6 (10.0, 27.6) | 5.4 (0.7, 39.2) |       |       |       |

**Menopausal status**

| Pre | 17.6 (14.0, 22.1) | 0.49 | 19.8 (15.7, 24.9) | 0.02 | NA | NA | NA |       |
| Peri| 17.8 (14.4, 22.0) | 16.4 (10.6, 25.4) |       |       |       |       |       |       |
| Post| 15.8 (12.5, 20.0) | 15.4 (12.1, 19.4) |       |       |       |       |       |       |

**Menopausal hormone therapy**

| None | 16.3 (13.2, 20.2) | 0.34 | 15.6 (12.6, 19.2) | 0.50 | 15.9 (13.9, 18.3) | 0.37 | 17.2 (15.1, 19.4) | 0.14 | 15.2 (12.8, 17.9) | 0.13 |
| Combination | 15.2 (8.2, 28.2) | 18.3 (12.7, 26.3) | 16.9 (12.6, 22.6) | 33.0 (23.7, 45.9) | 21.9 (14.0, 34.2) |       |       |       |
| Estrogen only | 26.4 (14.7, 47.2) | 18.3 (12.3, 27.2) | 21.0 (14.5, 30.2) | 25.3 (17.7, 36.1) | 32.2 (22.2, 46.4) |       |       |       |
Table 17. Rates of Recommendations for Biopsy After Screening With Digital Mammography by Risk Factors*

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral contraceptives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No current</td>
<td>16.7 (13.6, 20.6)</td>
<td>16.0 (13.1, 19.5)</td>
<td>0.32</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Current use</td>
<td>12.5 (9.5, 16.3)</td>
<td>13.0 (7.0, 24.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>21.4 (17.0, 26.8)</td>
<td>19.3 (14.7, 25.1)</td>
<td>0.40</td>
<td>17.4 (14.4, 21.0)</td>
<td>0.12</td>
</tr>
<tr>
<td>25 to &lt;30</td>
<td>17.6 (13.7, 22.6)</td>
<td>18.0 (13.3, 24.4)</td>
<td>18.9 (15.3, 23.4)</td>
<td>21.9 (18.2, 26.3)</td>
<td>16.6 (12.5, 21.9)</td>
</tr>
<tr>
<td>≥30</td>
<td>15.3 (12.3, 19.2)</td>
<td>18.4 (14.5, 23.4)</td>
<td>22.2 (18.1, 27.2)</td>
<td>26.7 (21.9, 32.4)</td>
<td>26.6 (18.5, 38.1)</td>
</tr>
</tbody>
</table>

*2-sided P-values and 95% confidence intervals from a logistic regression model that accounts for clustering by radiology facility using generalized estimating equations.
†Categories include: almost entirely fat=fat; scattered fibroglandular densities=scattered; heterogeneously dense=heterogeneous; and extremely dense=extreme.

**Abbreviations:** kg=kilogram; m=meter; NA=not applicable; peri=perimenopausal; pre=premenopausal; post=postmenopausal.
### Table 18. Rates of False-Positive and False-Negative Digital Mammography Results and Recommendations for Additional Imaging and Biopsies by Different Breast Density Categories*

<table>
<thead>
<tr>
<th></th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women screened, n</strong></td>
<td>113,770</td>
<td>127,958</td>
<td>94,507</td>
<td>50,204</td>
<td>18,752</td>
</tr>
<tr>
<td><strong>False-positive mammography results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number per 1,000 women screened per round (95% CI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A Fat-Scattered</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>108.4 (95.6, 122.7)</td>
<td>80.5 (71.1, 90.9)</td>
<td>74.1 (66.4, 82.6)</td>
<td>67.3 (60.4, 74.9)</td>
<td>60.3 (54.0, 67.4)</td>
</tr>
<tr>
<td>Extreme</td>
<td>112.1 (94.4, 132.7)</td>
<td>92.7 (77.5, 110.5)</td>
<td>75.2 (64.7, 87.1)</td>
<td>57.7 (43.9, 75.5)</td>
<td>85.1 (61.7, 116.2)</td>
</tr>
<tr>
<td><strong>B Fat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scattered</td>
<td>116.8 (102.9, 132.3)</td>
<td>87.7 (77.1, 99.6)</td>
<td>81.6 (72.7, 91.4)</td>
<td>73.4 (65.4, 82.2)</td>
<td>65.8 (58.4, 73.9)</td>
</tr>
<tr>
<td>Heterogeneous-Extreme</td>
<td>135.3 (113.9, 160.0)</td>
<td>112.0 (96.9, 129.2)</td>
<td>96.9 (88.4, 110.4)</td>
<td>86.2 (76.4, 97.1)</td>
<td>82.7 (72.6, 93.9)</td>
</tr>
<tr>
<td><strong>C Fat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scattered</td>
<td>116.8 (102.9, 132.3)</td>
<td>87.7 (77.1, 99.6)</td>
<td>81.6 (72.7, 91.4)</td>
<td>73.4 (65.4, 82.2)</td>
<td>65.8 (58.4, 73.9)</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>142.2 (120.2, 167.4)</td>
<td>115.8 (100.3, 133.2)</td>
<td>101.8 (91.0, 113.8)</td>
<td>88.7 (78.7, 99.9)</td>
<td>82.4 (72.6, 93.5)</td>
</tr>
<tr>
<td>Extreme</td>
<td>112.1 (94.4, 132.7)</td>
<td>92.7 (77.5, 110.5)</td>
<td>75.2 (64.7, 87.1)</td>
<td>57.7 (43.9, 75.5)</td>
<td>85.1 (61.7, 116.2)</td>
</tr>
<tr>
<td><strong>D Fat-Scattered</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneous-Extreme</td>
<td>135.3 (113.9, 160.0)</td>
<td>112.0 (96.9, 129.2)</td>
<td>98.9 (88.4, 110.4)</td>
<td>86.2 (76.4, 97.1)</td>
<td>82.7 (72.6, 93.9)</td>
</tr>
<tr>
<td><strong>False-negative mammography results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number per 1,000 women screened per round (95% CI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A Fat-Scattered</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>0.4 (0.3, 0.6)</td>
<td>0.6 (0.4, 0.8)</td>
<td>0.002</td>
<td>0.8 (0.5, 1.1)</td>
<td>0.006</td>
</tr>
<tr>
<td>Extreme</td>
<td>1.7 (1.2, 2.5)</td>
<td>1.6 (0.9, 2.8)</td>
<td>1.2 (0.6, 2.7)</td>
<td>5.6 (2.4, 12.9)</td>
<td>6.9 (2.5, 18.5)</td>
</tr>
<tr>
<td><strong>B Fat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scattered</td>
<td>0.2 (0.0, 0.9)</td>
<td>0.3 (0.1, 0.7)</td>
<td>0.001</td>
<td>0.6 (0.2, 1.5)</td>
<td>0.007</td>
</tr>
<tr>
<td>Heterogeneous-Extreme</td>
<td>1.4 (1.0, 1.7)</td>
<td>1.5 (1.1, 1.9)</td>
<td>1.6 (1.2, 2.2)</td>
<td>2.6 (1.8, 3.7)</td>
<td>1.7 (0.8, 3.3)</td>
</tr>
<tr>
<td><strong>C Fat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scattered</td>
<td>0.2 (0.0, 0.9)</td>
<td>0.3 (0.1, 0.7)</td>
<td>0.001</td>
<td>0.6 (0.2, 1.5)</td>
<td>0.02</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>1.3 (1.0, 1.7)</td>
<td>1.4 (1.0, 2.0)</td>
<td>1.7 (1.3, 2.3)</td>
<td>2.3 (1.6, 3.4)</td>
<td>1.1 (0.5, 2.4)</td>
</tr>
<tr>
<td>Extreme</td>
<td>1.7 (1.2, 2.5)</td>
<td>1.6 (0.9, 2.8)</td>
<td>1.2 (0.6, 2.7)</td>
<td>5.6 (2.4, 12.9)</td>
<td>6.9 (2.5, 18.5)</td>
</tr>
<tr>
<td><strong>D Fat-Scattered</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneous-Extreme</td>
<td>1.4 (1.0, 1.7)</td>
<td>1.5 (1.1, 1.9)</td>
<td>1.6 (1.2, 2.2)</td>
<td>2.6 (1.8, 3.7)</td>
<td>1.7 (0.8, 3.3)</td>
</tr>
</tbody>
</table>
Table 18. Rates of False-Positive and False-Negative Digital Mammography Results and Recommendations for Additional Imaging and Biopsies by Different Breast Density Categories*

<table>
<thead>
<tr>
<th>Age, y</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recommendations for additional imaging</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number per 1,000 women screened per round (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>Fat-Scattered</td>
<td>110.8 (97.9, 125.2)</td>
<td>84.4 (74.8, 95.1)</td>
<td>&lt;0.001</td>
<td>81.0 (73.1, 89.6)</td>
</tr>
<tr>
<td></td>
<td>Heterogeneous</td>
<td>146.0 (123.9, 171.3)</td>
<td>121.6 (105.8, 139.3)</td>
<td>110.6 (97.9, 122.6)</td>
<td>98.0 (87.9, 111.4)</td>
</tr>
<tr>
<td></td>
<td>Extreme</td>
<td>116.5 (98.4, 137.4)</td>
<td>98.4 (83.1, 116.2)</td>
<td>81.0 (70.3, 93.2)</td>
<td>63.3 (49.7, 80.1)</td>
</tr>
<tr>
<td>B</td>
<td>Fat</td>
<td>64.4 (52.3, 79.1)</td>
<td>54.1 (46.5, 62.8)</td>
<td>&lt;0.001</td>
<td>53.4 (47.9, 59.4)</td>
</tr>
<tr>
<td></td>
<td>Scattered</td>
<td>119.4 (105.5, 135.0)</td>
<td>92.1 (81.2, 104.3)</td>
<td>89.0 (79.9, 99.1)</td>
<td>82.1 (73.9, 91.2)</td>
</tr>
<tr>
<td></td>
<td>Heterogeneous-Extreme</td>
<td>139.3 (117.7, 164.1)</td>
<td>117.8 (102.4, 135.3)</td>
<td>107.3 (96.7, 118.9)</td>
<td>96.1 (85.5, 107.9)</td>
</tr>
<tr>
<td>C</td>
<td>Fat</td>
<td>64.4 (52.3, 79.1)</td>
<td>54.1 (46.5, 62.8)</td>
<td>&lt;0.001</td>
<td>53.4 (47.9, 59.4)</td>
</tr>
<tr>
<td></td>
<td>Scattered</td>
<td>119.4 (105.5, 135.0)</td>
<td>92.1 (81.2, 104.3)</td>
<td>89.0 (79.9, 99.1)</td>
<td>82.1 (73.9, 91.2)</td>
</tr>
<tr>
<td></td>
<td>Heterogeneous</td>
<td>146.0 (123.9, 171.3)</td>
<td>121.6 (105.8, 139.3)</td>
<td>110.6 (99.7, 122.6)</td>
<td>99.0 (87.9, 111.4)</td>
</tr>
<tr>
<td></td>
<td>Extreme</td>
<td>116.5 (98.4, 137.4)</td>
<td>98.4 (83.1, 116.2)</td>
<td>81.0 (70.3, 93.2)</td>
<td>63.3 (49.7, 80.1)</td>
</tr>
<tr>
<td>D</td>
<td>Fat-Scattered</td>
<td>110.8 (97.9, 125.2)</td>
<td>84.4 (74.8, 95.1)</td>
<td>&lt;0.001</td>
<td>81.0 (73.1, 89.6)</td>
</tr>
<tr>
<td></td>
<td>Heterogeneous-Extreme</td>
<td>139.3 (117.7, 164.1)</td>
<td>117.8 (102.4, 135.3)</td>
<td>107.3 (96.7, 118.9)</td>
<td>96.1 (85.5, 107.9)</td>
</tr>
<tr>
<td><strong>Recommendations for biopsy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number per 1,000 women screened per round (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>Fat-Scattered</td>
<td>12.2 (9.9, 15.0)</td>
<td>11.8 (9.6, 14.5)</td>
<td>&lt;0.001</td>
<td>15.6 (13.7, 17.7)</td>
</tr>
<tr>
<td></td>
<td>Heterogeneous</td>
<td>18.9 (15.8, 22.5)</td>
<td>20.2 (17.3, 23.7)</td>
<td>19.3 (16.9, 22.2)</td>
<td>21.0 (18.0, 24.5)</td>
</tr>
<tr>
<td></td>
<td>Extreme</td>
<td>20.2 (16.8, 24.3)</td>
<td>19.2 (14.3, 25.7)</td>
<td>13.8 (10.5, 18.2)</td>
<td>13.0 (7.2, 23.3)</td>
</tr>
<tr>
<td>B</td>
<td>Fat</td>
<td>7.5 (5.5, 10.1)</td>
<td>8.4 (6.0, 11.7)</td>
<td>&lt;0.001</td>
<td>11.7 (9.5, 14.6)</td>
</tr>
<tr>
<td></td>
<td>Scattered</td>
<td>13.1 (10.6, 16.1)</td>
<td>12.7 (10.3, 15.6)</td>
<td>16.7 (14.7, 19.0)</td>
<td>17.1 (14.9, 19.6)</td>
</tr>
<tr>
<td></td>
<td>Heterogeneous-Extreme</td>
<td>19.2 (16.2, 22.7)</td>
<td>20.1 (16.9, 23.7)</td>
<td>18.7 (16.6, 21.2)</td>
<td>20.4 (17.4, 23.8)</td>
</tr>
</tbody>
</table>
### Table 18. Rates of False-Positive and False-Negative Digital Mammography Results and Recommendations for Additional Imaging and Biopsies by Different Breast Density Categories*

<table>
<thead>
<tr>
<th>Age, y</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat</td>
<td>7.5</td>
<td>(5.5, 10.1)</td>
<td>&lt;0.001</td>
<td>8.4</td>
<td>(6.0, 11.7)</td>
</tr>
<tr>
<td>Scattered</td>
<td>13.1</td>
<td>(10.6, 16.1)</td>
<td>12.7</td>
<td>(10.3, 15.6)</td>
<td>16.7</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>18.9</td>
<td>(15.8, 22.5)</td>
<td>20.2</td>
<td>(17.3, 23.7)</td>
<td>19.3</td>
</tr>
<tr>
<td>Extreme</td>
<td>20.2</td>
<td>(16.8, 24.3)</td>
<td>19.2</td>
<td>(14.3, 25.7)</td>
<td>13.8</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat-Scattered</td>
<td>12.2</td>
<td>(9.9, 15.0)</td>
<td>&lt;0.001</td>
<td>11.8</td>
<td>(9.6, 14.5)</td>
</tr>
<tr>
<td>Heterogeneous-Extreme</td>
<td>19.2</td>
<td>(16.2, 22.7)</td>
<td>20.1</td>
<td>(16.9, 23.7)</td>
<td>18.7</td>
</tr>
</tbody>
</table>

*2-sided P-values and 95% confidence intervals from a logistic regression model that accounts for clustering by radiology facility using generalized estimating equations.
†Categories include: almost entirely fat=fat; scattered fibroglandular densities=scattered; heterogeneously dense=heterogeneous; and extremely dense=extreme.
Table 19. U.S. Studies of Cumulative False-Positive Mammography and Biopsy Results

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Population; age, years; participants, n</th>
<th>Study years; comparison</th>
<th>Outcome measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>New studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hubbard et al, 2011</td>
<td>Post-intervention</td>
<td>U.S., 7 mammography registries in the BCSC; 40-59; 169,456</td>
<td>1994-2006; annual vs. biennial screening by age</td>
<td>FP results (no diagnosis of invasive carcinoma or DCIS within 1 year of screening or before the next screening); recalls (BIRADS 0, 3, 4, 5)</td>
<td>Cumulative probability of FP mammography after 10 years, % (95% CI)</td>
</tr>
<tr>
<td></td>
<td>series</td>
<td></td>
<td></td>
<td></td>
<td>• Age 40: annual, 61.3 (59.4 to 63.1); biennial, 41.6 (40.6 to 42.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Age 50: annual, 61.3 (58.0 to 64.7); biennial, 42.0 (40.4 to 43.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cumulative probability of FP biopsy after 10 years, % (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Age 40: annual, 7.0 (6.1 to 7.8); biennial, 4.8 (4.4 to 5.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Age 50: annual, 9.4 (7.4 to 11.5); biennial, 6.4 (5.6 to 7.2)</td>
</tr>
<tr>
<td>Kerlikowske et al,</td>
<td>Post-intervention</td>
<td>U.S., 7 mammography registries in the BCSC; 40-74; 11,474 with breast cancer, 922,2624 without</td>
<td>1994-2008; annual vs. biennial vs. triennial screening by age, breast density, and menopausal hormone therapy</td>
<td>FP results (no diagnosis of invasive carcinoma or DCIS within 1 year of screening or before the next screening mammogram); recalls (BIRADS 0, 3, 4, 5)</td>
<td>Cumulative probability of FP mammography after 10 years, by breast density,* % (95% CI)</td>
</tr>
<tr>
<td>2013</td>
<td>series</td>
<td></td>
<td></td>
<td></td>
<td>• 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).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• 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).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cumulative probability of FP biopsy after 10 years, by breast density,* % (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• 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).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• 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).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>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 HT.</td>
</tr>
<tr>
<td>2009 review</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elmore et al, 1998</td>
<td>Post-intervention</td>
<td>U.S., randomly sampled patients from 11 health centers in an HMO; 40-69</td>
<td>1983-1995; annual vs. biennial screening</td>
<td>FP test results (not a true positive=breast cancer diagnosed on the basis of pathological findings within 1 year of mammography)</td>
<td>Cumulative risk of at least one FP after 10 screening mammograms, % (95% CI)</td>
</tr>
<tr>
<td></td>
<td>series</td>
<td></td>
<td></td>
<td></td>
<td>• 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).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cumulative risk of FP biopsy, % (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Overall: 19 (9.8 to 41.2).</td>
</tr>
</tbody>
</table>

*Categories include: fatty; scattered fibroglandular densities; heterogeneously dense; extremely dense.

**Abbreviations:** BCSC=Breast Cancer Surveillance Consortium; BIRADS=Breast Imaging Reporting and Data System; CI=confidence interval; DCIS=ductal carcinoma in situ; FP=false-positive; HMO=health maintenance organization; HT=hormone therapy; n=number; U.S.=United States; vs.=versus.
## Table 20. Studies of Overdiagnosis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Age, years</th>
<th>Study years</th>
<th>Data source</th>
<th>Comparison groups</th>
<th>Approach, lead time adjustment</th>
<th>Overdiagnosis measures as defined by each study</th>
<th>Overdiagnosis rate of invasive cancer + DCIS</th>
<th>Overdiagnosis rate of invasive cancer</th>
<th>Overdiagnosis rate of DCIS</th>
</tr>
</thead>
</table>
| Bleyer and Welch, 2012 | ≥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 | 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: 7.3%  
  • Population estimate: 6.7%  | Participation estimate: 5.4%  
  • Population estimate: -0.7%  | NR         | NR          |
Table 20. Studies of Overdiagnosis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Age, years</th>
<th>Study years</th>
<th>Data source</th>
<th>Comparison groups</th>
<th>Approach, lead time adjustment</th>
<th>Overdiagnosis measures as defined by each study</th>
<th>Overdiagnosis rate of invasive cancer + DCIS</th>
<th>Overdiagnosis rate of invasive cancer</th>
<th>Overdiagnosis rate of DCIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>de Gelder et al, 2011&lt;sup&gt;187&lt;/sup&gt;</td>
<td>49-74</td>
<td>2004-2006</td>
<td>Screening program (biennial); Netherlands</td>
<td>Modeled incidence of screening vs. predicted incidence without screening</td>
<td>LT; statistical adjustment; preclinical DCIS: mean 5.2 years; preclinical invasive: 2.6 years</td>
<td>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.  • Non-progressive model: no preclinical screen-detected DCIS, majority regress, 2% are clinically diagnosed.</td>
<td>• Baseline model: 2.5% all cases; 8.2% screen-detected  • Progressive model: 1.4% all cases; 5.0% screen-detected  • Non-progressive model: 7.7% all cases; 25.2% screen-detected</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>de Gelder et al, 2011&lt;sup&gt;188, 189&lt;/sup&gt;</td>
<td>0-69; 0-74</td>
<td>1990-1998; 1998-2007</td>
<td>Screening program (biennial); Netherlands</td>
<td>Modeled incidence of screening vs. predicted incidence without screening</td>
<td>LT; compensatory drop; mean 2.6 years</td>
<td>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 years.</td>
<td>• 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 years.</td>
<td>Overall: 4%-7%  • Swedish Trial: 4.3 cases per 1,000 women screened for 20 years  • U.K. Program: 2.3 cases per 1,000 women screened for 20 years</td>
<td>NR</td>
</tr>
<tr>
<td>Duffy et al, 2010&lt;sup&gt;110&lt;/sup&gt;</td>
<td>50-69</td>
<td>1977-1998; 1974-2003</td>
<td>Swedish Two-County Trial; U.K. National Breast Screening Program</td>
<td>Active vs. passive screening; population before vs. after widespread screening</td>
<td>EI; compensatory drop</td>
<td>• 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 years.</td>
<td>• Overall: 4%-7%  • Swedish Trial: 4.3 cases per 1,000 women screened for 20 years  • U.K. Program: 2.3 cases per 1,000 women screened for 20 years</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>
Table 20. Studies of Overdiagnosis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Age, years</th>
<th>Study years</th>
<th>Data source</th>
<th>Comparison groups</th>
<th>Approach, lead time adjustment</th>
<th>Overdiagnosis measures as defined by each study</th>
<th>Overdiagnosis rate of invasive cancer + DCIS</th>
<th>Overdiagnosis rate of invasive cancer</th>
<th>Overdiagnosis rate of DCIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Falk et al, 2013</td>
<td>50-69</td>
<td>1995-2009</td>
<td>Norwegian Breast Cancer Screening Program (biennial)</td>
<td>Women screened vs. those never invited or did not attend screening</td>
<td>EI; compensatory drop</td>
<td>Women attending screening adjusted for compliance with screening vs. 3 reference rates:</td>
<td>16.5%; 16.3%; 13.9%</td>
<td>11.3%; 11.2%; 9.6%</td>
<td>NR</td>
</tr>
<tr>
<td>Gunsoy et al, 2014</td>
<td>40-73</td>
<td>1971-2010</td>
<td>Data from various sources in the U.K.</td>
<td>Women screened vs. not screened</td>
<td>Multiple statistical adjustments</td>
<td>Markov model of the difference between cumulative incidence of invasive + DCIS with denominators:</td>
<td>All cases: 4.3 to 8.9%</td>
<td>Screening period: 6.7 to 10.1%</td>
<td>NR</td>
</tr>
<tr>
<td>Hellquist et al, 2012</td>
<td>40-49</td>
<td>1986-2005</td>
<td>Screening for Young Women Trial; Sweden</td>
<td>Population in areas with vs. without screening</td>
<td>EI; statistical adjustment; up to 1.5 years</td>
<td>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).</td>
<td>Rate ratio: 1.01 (95% CI 0.94 to 1.08)</td>
<td>Rate ratio: 0.95 (95% CI 0.88 to 1.01)</td>
<td>NR</td>
</tr>
</tbody>
</table>
### Table 20. Studies of Overdiagnosis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Age, years</th>
<th>Study years</th>
<th>Data source</th>
<th>Comparison groups</th>
<th>Approach, lead time adjustment</th>
<th>Overdiagnosis measures as defined by each study</th>
<th>Overdiagnosis rate of invasive cancer + DCIS</th>
<th>Overdiagnosis rate of invasive cancer</th>
<th>Overdiagnosis rate of DCIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kalager et al, 2012</td>
<td>50-69</td>
<td>1996-2005</td>
<td>Norwegian Breast Cancer Screening Program (biennial)</td>
<td>Population in areas with vs. without screening</td>
<td>EI; compensatory drop; Approach 1: 10-year lead time; Approach 2: 5 or 2-year</td>
<td>• Approach 1: Incidence rates in the screening and non-screening groups for women aged 50-79 years. • Approach 2: Excluded all cases of cancer detected in the first screening round, compares incidence in screened women vs. women 2-5 years older.</td>
<td>NR</td>
<td>• Approach 1: entire country: 25%, region 1: 18% • Approach 2: 5-year lead time: 15%, 2-year lead time: 20%</td>
<td>NR</td>
</tr>
<tr>
<td>Marmot et al, 2013</td>
<td>40-69</td>
<td>1976-2001</td>
<td>Meta-analysis of Canadian National Breast Screening Study and Malmö I Trial</td>
<td>Randomized trials; screening vs. usual care</td>
<td>EI; none</td>
<td>Excess incidence of breast cancer (both invasive cancer and DCIS) in the screening population was compared with the incidence in the absence of screening</td>
<td>NR</td>
<td>Short case accrual: 19.0% (95% CI, 15.2-22.7%; 3 trials) • Long case accrual: 10.7% (95% CI, 9.3-12.2%; 3 trials)</td>
<td>NR</td>
</tr>
<tr>
<td>Martinez-Alonso et al, 2010</td>
<td>40-69</td>
<td>1980-2004</td>
<td>Cancer registry; Catalonia, Spain</td>
<td>Modeled pre vs. post screening incidence</td>
<td>EI; statistical adjustments</td>
<td>Probabilistic model for birth cohorts: 1935, 1940, 1945, 1950; observed vs. expected cumulative incidence.</td>
<td>NR</td>
<td>1935: 0.4% 1940: 23.3% 1945: 30.6% 1950: 46.6%</td>
<td>NR</td>
</tr>
<tr>
<td>Miller et al, 2014</td>
<td>40-59</td>
<td>1980-1985</td>
<td>Canadian National Breast Screening Study</td>
<td>Randomized trial; screening vs. usual care</td>
<td>EI; none</td>
<td>Excess of breast cancer cases in mammography arm vs. the control arm of trial.</td>
<td>NR</td>
<td>22% of screen-detected cancer</td>
<td>NR</td>
</tr>
</tbody>
</table>
Table 20. Studies of Overdiagnosis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Age, years</th>
<th>Study years</th>
<th>Data source</th>
<th>Comparison groups</th>
<th>Approach, lead time adjustment</th>
<th>Overdiagnosis measures as defined by each study</th>
<th>Overdiagnosis rate of invasive cancer + DCIS</th>
<th>Overdiagnosis rate of invasive cancer</th>
<th>Overdiagnosis rate of DCIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morrell et al, 2010&lt;sup&gt;74&lt;/sup&gt;</td>
<td>50-69</td>
<td>1999-2001</td>
<td>Screening program (biennial); Australia</td>
<td>Screened vs. unscreened age group or prior to screening implementation</td>
<td>EI; statistical adjustment; 2- or 5-year lead times</td>
<td>Observed annual incidence minus expected annual incidence divided by expected annual incidence; • Interpolation approach: incidence in unscreened women (≤40 or ≥80) modeled by 5-year age group. • Extrapolation approach: incidence for the period prior to the introduction of screening modeled for all 5-year age groups and extrapolated to 1999-2001.</td>
<td>NR</td>
<td>• Interpolation: 2-year: 51%; 5-year: 42%; • Extrapolation: 2-year: 36%, 5-year: 30% • Rates higher for 50-59 vs. 60-69</td>
<td>NR</td>
</tr>
<tr>
<td>Njor et al, 2013&lt;sup&gt;75&lt;/sup&gt;</td>
<td>56-70</td>
<td>1991-2005</td>
<td>Screening program; Copenhagen and Funen, Denmark</td>
<td>Population in areas with vs. without screening</td>
<td>EI; compensatory drop</td>
<td>Cumulative incidence in screened population vs. expected incidence in unscreened counties</td>
<td>≥8 years followup: Copenhagen, 3% (-14% to 25%), Funen, 0.7% (-9% to 12%)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Puliti, et al, 2009&lt;sup&gt;76&lt;/sup&gt;</td>
<td>60-69</td>
<td>1990-NR</td>
<td>Screening program; Florence, Italy</td>
<td>Screening vs. pre-screening</td>
<td>EI; compensatory drop</td>
<td>Ratio of cumulative incidence of breast cancer in the invited group to those in the non-invited group at least 5 years after last screening, assuming 1.2% annual trend in pre-screening incidence.</td>
<td>Rate ratio: 1.01 (95% CI 0.95 to 1.07)</td>
<td>Rate ratio: 0.99 (95% CI 0.94 to 1.05)</td>
<td>NR</td>
</tr>
<tr>
<td>Seigneurin et al, 2011&lt;sup&gt;77&lt;/sup&gt;</td>
<td>50-69</td>
<td>1991-2006</td>
<td>Cancer registry; Isere, France</td>
<td>Modeled screening incidence</td>
<td>LT; statistical adjustment, 2-4 years</td>
<td>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.</td>
<td>NR</td>
<td>All diagnosed cancers: 1.5%, screen detected: 3.3%</td>
<td>All diagnosed cancers: 28.0%, screen detected: 31.9%</td>
</tr>
</tbody>
</table>
### Table 20. Studies of Overdiagnosis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Age, years</th>
<th>Study years</th>
<th>Data source</th>
<th>Comparison groups</th>
<th>Approach, lead time adjustment</th>
<th>Overdiagnosis measures as defined by each study</th>
<th>Overdiagnosis rate of invasive cancer + DCIS</th>
<th>Overdiagnosis rate of invasive cancer</th>
<th>Overdiagnosis rate of DCIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yen et al, 2012</td>
<td>40-74</td>
<td>1977-2005</td>
<td>Swedish Two County Trial; data from 1 county only (Dalarna)</td>
<td>Active screening vs. passive screening</td>
<td>EI; compensatory drop</td>
<td>Cumulative incidence in active screening vs. usual care groups</td>
<td>Relative risk: 1.00 (95% CI 0.92 to 1.08)</td>
<td>Relative risk: 0.99 (95% CI 0.88 to 1.55)</td>
<td>Relative risk: 1.17 (95% CI 0.88 to 1.55)</td>
</tr>
<tr>
<td>Zahl et al, 2012</td>
<td>40-79</td>
<td>1991-2009</td>
<td>Norway Cancer Registry</td>
<td>Screening vs. post-screening</td>
<td>EI; compensatory drop</td>
<td>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.</td>
<td>~50%</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>2009 report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Koning et al, 2006</td>
<td>50-74</td>
<td>1989-2001</td>
<td>National data from The Netherlands</td>
<td>Screening vs nonscreening (biennial)</td>
<td>Statistical adjustments; assumptions of DCIS progression</td>
<td>Microsimulation model</td>
<td>3% in screened population</td>
<td>8% screen-detected</td>
<td>NR</td>
</tr>
<tr>
<td>Duffy et al, 2005</td>
<td>40-74</td>
<td>1977-1985</td>
<td>Swedish Two-county Trial</td>
<td>Active vs. passive screening</td>
<td>Lead time statistical adjustments</td>
<td>Markov multistate model</td>
<td>1% in screened population</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td>39-59</td>
<td>1982-1996</td>
<td>Gothenburg trial</td>
<td>Screening vs. no screening</td>
<td>Lead time statistical adjustments</td>
<td>Markov multistate model</td>
<td>2% in screened population</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Olsen et al, 2006</td>
<td>50-71</td>
<td>1991-1996</td>
<td>Copenhagen, Denmark screening program (biennial)</td>
<td>Incidence in screened women</td>
<td>Statistical adjustments</td>
<td>Chronic disease statistical model of screen-detected overdiagnosis</td>
<td>Prevalence: 7.8%; Incidence: 0.5%</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Paci et al, 2004</td>
<td>50-69</td>
<td>1985-1999</td>
<td>Florence, Italy; screening program</td>
<td>Incidence in screening vs. prescreening</td>
<td>EI; corrected for lead time</td>
<td>Observed/expected cases</td>
<td>5%</td>
<td>2%</td>
<td>3%</td>
</tr>
<tr>
<td>Paci et al, 2006</td>
<td>50-74</td>
<td>1986-2001</td>
<td>Italy; screening program</td>
<td>Prescreening incidence</td>
<td>EI; corrected for lead time</td>
<td>Observed/expected cases</td>
<td>4.6%; range -0.6% to 9.7% varies by age (highest in 50-54 and 65-74)</td>
<td>3.2%</td>
<td>1.4%</td>
</tr>
<tr>
<td>Author, year</td>
<td>Age, years</td>
<td>Study years</td>
<td>Data source</td>
<td>Comparison groups</td>
<td>Approach, lead time adjustment</td>
<td>Overdiagnosis measures as defined by each study</td>
<td>Overdiagnosis rate of invasive cancer + DCIS</td>
<td>Overdiagnosis rate of invasive cancer</td>
<td>Overdiagnosis rate of DCIS</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>-------------</td>
<td>-------------</td>
<td>------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------</td>
<td>------------------------------------------</td>
<td>----------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Yen et al, 2003</td>
<td>40-69</td>
<td>Screening vs. not screening</td>
<td>LT; statistical adjustment</td>
<td>Six state Markov model</td>
<td>NR</td>
<td>NR</td>
<td>Prevalence: 37% Incidence: 4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40-69</td>
<td>Screening vs. not screening</td>
<td>LT; statistical adjustment</td>
<td>Six state Markov model</td>
<td>NR</td>
<td>NR</td>
<td>40-49: 19%, 3% 50-59: 23%, 4% 60-69: 46%, 6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zackrisson et al, 2006</td>
<td>55-69</td>
<td>Randomized screening vs. no screening</td>
<td>EI; compensatory drop</td>
<td>Comparison of incidence in screened vs. unscreened</td>
<td>10% of incidence in control group</td>
<td>7%</td>
<td>3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zahl et al, 2004</td>
<td>50-69</td>
<td>Prescreening incidence</td>
<td>EI; compensatory drop</td>
<td>Changes in age-specific incidence rates associated with the introduction of screening programs</td>
<td>NR</td>
<td>30% of incidence in screened population</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*An additional 6 model estimates for each year are published in this paper to show the range of estimates varies by selection of the denominator.
†Population overlap with Kalager; may have data to calculate by age group.
‡Same Copenhagen population as Olsen.
§Population overlap with Falk.

**Abbreviations:** CI=confidence interval; DCIS=ductal carcinoma in situ; EI=excess incidence approach; LT=lead time approach; NR=not reported; SEER=Surveillance, Epidemiology, and End Results Program; U.K.=United Kingdom; vs.=versus.
Table 21. Overdiagnosis Estimates From Randomized, Controlled Trials Without Screening of Control Groups

<table>
<thead>
<tr>
<th>Trial (reference)</th>
<th>Age, years</th>
<th>Overdiagnosis, % (95% CI)</th>
<th>Overdiagnosis, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malmö I&lt;sup&gt;5&lt;/sup&gt;</td>
<td>55-69</td>
<td>18.7 (15.1 to 22.4)</td>
<td>10.5 (8.4 to 12.7)</td>
</tr>
<tr>
<td>CNBSS-I&lt;sup&gt;6&lt;/sup&gt;</td>
<td>40-49</td>
<td>22.7 (18.4 to 27.0)</td>
<td>12.4 (9.9 to 14.9)</td>
</tr>
<tr>
<td>CNBSS-2&lt;sup&gt;6&lt;/sup&gt;</td>
<td>50-59</td>
<td>16.0 (12.5 to 19.5)</td>
<td>9.7 (7.5 to 11.9)</td>
</tr>
<tr>
<td>Meta-analysis&lt;sup&gt;6,7&lt;/sup&gt;</td>
<td>40-69</td>
<td>19.0 (15.2 to 22.7; ( I^2 = 64.8% ); ( p = 0.058 ))</td>
<td>10.7 (9.3 to 12.2; ( I^2 = 22.3% ); ( p = 0.276 ))</td>
</tr>
</tbody>
</table>

*Excess cancers as a proportion of cancers diagnosed during the screening period (short-case accrual) or over the followup period (long-case accrual) in women invited for screening.

**Abbreviations**: CI=confidence interval; CNBSS=Canadian National Breast Screening Study.
Table 22. Systematic Reviews of Psychological Harms From False-Positive vs. Normal Screening Mammography Results

<table>
<thead>
<tr>
<th>Author, year quality rating</th>
<th>Inclusion criteria</th>
<th>Searches</th>
<th>Number of studies; number of participants</th>
<th>Re-attendance</th>
<th>Anxiety</th>
<th>Depression</th>
<th>Breast cancer worry/distress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond et al, 2013&lt;sup&gt;185&lt;/sup&gt; Good</td>
<td>Studies in the U.K. comparing psychological and behavioral outcomes of women with FP vs. normal screening mammograms.</td>
<td>Multiple databases through November 2011</td>
<td>7 studies* of psychological harms N=3,168; re-attendance N=151,490</td>
<td>• Lower in FP vs. normal (2 studies)</td>
<td>No difference (2 studies)</td>
<td>No difference (2 studies)</td>
<td>Higher in FP vs. normal (3 studies)</td>
</tr>
<tr>
<td>Hafslund and Nortvedt, 2009&lt;sup&gt;194&lt;/sup&gt; Fair</td>
<td>Studies of women not at high risk; ages 40-74 years invited to mammography screening.</td>
<td>Multiple databases January 1995 to July 2007</td>
<td>17 studies† N=18,097</td>
<td>NR</td>
<td>Higher in FP vs. normal (15 studies)</td>
<td>NR</td>
<td>Higher in FP vs. normal (15 studies)</td>
</tr>
<tr>
<td>2009 Review</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brewer et al, 2007&lt;sup&gt;184&lt;/sup&gt; Fair</td>
<td>Studies comparing psychological and behavioral outcomes of women with FP vs. normal screening mammograms.</td>
<td>Multiple databases through September 2006</td>
<td>23 studies N=313,967</td>
<td>• Lower in FP vs. normal in U.S. (RR 1.07; 95% CI 1.02 to 1.12; 5 studies)</td>
<td>• Higher in FP vs. normal (4 studies)</td>
<td>• Lower in FP vs. normal (7 studies)</td>
<td>• Higher in FP vs. normal (4 studies)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Lower in normal vs. FP in Canada (RR 0.63; 95% CI 0.50 to 0.80; 2 studies)</td>
<td>• No differences (4 studies)</td>
<td>• No differences (1 study)</td>
<td>• No differences (3 studies)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• No differences in Europe (RR 0.97; 95% CI 0.93 to 1.01; 5 studies)</td>
<td>• Conflicting results over time (2 studies)</td>
<td>• Conflicting results based on measure (1 study)</td>
<td>• Conflicting results over time (2 studies)</td>
</tr>
<tr>
<td>Brett et al, 2005&lt;sup&gt;183&lt;/sup&gt; Fair</td>
<td>Studies of the psychological impact of mammography screening.</td>
<td>Multiple databases 1982 to 2003</td>
<td>54 studies N=NR</td>
<td>NR</td>
<td>Higher in FP vs. normal (14 studies)</td>
<td>NR</td>
<td>Higher in FP vs. normal (9 studies)</td>
</tr>
</tbody>
</table>

*5 of 7 studies were included in at least one of the systematic reviews included in the 2009 review.
†13 of 17 studies were included in at least one of the systematic reviews included in the 2009 review.

**Abbreviations:** CI=confidence interval; FP=false-positive; N=number; NR=not reported; RR=risk ratio; U.K.=United Kingdom; U.S.=United States; vs.=versus.
Table 23. Summary of Results of New Studies of Psychological Harms

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Population</th>
<th>Comparisons, N</th>
<th>Measures</th>
<th>Re-attendance</th>
<th>Anxiety</th>
<th>Depression</th>
<th>Breast cancer worry</th>
<th>General QOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bredal et al, 2013</td>
<td>Before-after</td>
<td>Women recalled in a screening program in Norway</td>
<td>A) At recall (n=640) B) 4 weeks later</td>
<td>HADS (score ≥11)</td>
<td>NR</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Brodersen and Siersma, 2013</td>
<td>Nested case-control</td>
<td>Screening programs in Denmark</td>
<td>A) FP (n=272) B) Normal (n=864) C) TP (n=174)</td>
<td>COS-BC</td>
<td>NR</td>
<td>• Immediate: higher A+C vs. B; no difference A vs. C • 3-years after: higher C vs. A+B and A vs. B</td>
<td>NR</td>
<td>0</td>
<td>NR</td>
</tr>
<tr>
<td>Espasa et al, 2012</td>
<td>Case-control</td>
<td>Screening program in Spain</td>
<td>A) FP (n=100) B) Normal (n=50)</td>
<td>HADS, structured interview</td>
<td>NR</td>
<td>0</td>
<td>0</td>
<td>Higher FP vs. normal</td>
<td>NR</td>
</tr>
<tr>
<td>Fitzpatrick et al, 2011</td>
<td>Retrospective cohort</td>
<td>Screening program in the United Kingdom</td>
<td>A) FP (n=9,746) B) Normal (n=148,589)</td>
<td>Re-attendance</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Gibson et al, 2009</td>
<td>Prospective cohort</td>
<td>New Hampshire Mammography Network and New Hampshire Women for Health study</td>
<td>A) FP (n=2,107) B) Normal (n=11,384)</td>
<td>WHQ</td>
<td>NR</td>
<td>Higher for non-white with FP vs. normal</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Hafslund et al, 2012</td>
<td>Nested case-control</td>
<td>Screening programs in Norway</td>
<td>A) FP (n=128) B) Normal (n=195)</td>
<td>SF-36, HADS</td>
<td>NR</td>
<td>0</td>
<td>More cases for FP vs. normal</td>
<td>NR</td>
<td>Lower for FP vs. normal</td>
</tr>
<tr>
<td>Keyzer-Dekker, 2012</td>
<td>Prospective cohort</td>
<td>Women with abnormal results in The Netherlands</td>
<td>A) 1st screen recalls (n=186) B) Repeat screen recalls (n=296)</td>
<td>STAI, NEO-FFI, CES-D, WHOQOL</td>
<td>NR</td>
<td>0*</td>
<td>0*</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Klompenhouwer et al, 2014</td>
<td>Retrospective cohort</td>
<td>Screening program in The Netherlands</td>
<td>A) Normal screen (n=373,474) B) 1st screen recalls (n=6,672) C) repeat screen recalls for different lesion (n=161) D) repeat screen</td>
<td>Re-attendance</td>
<td>NR</td>
<td>93.2% vs. 65.4% vs. 56.7% vs. 44.3%</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Author, year</td>
<td>Study design</td>
<td>Population</td>
<td>Comparisons, N</td>
<td>Measures</td>
<td>Re-attendance</td>
<td>Anxiety</td>
<td>Depression</td>
<td>Breast cancer worry</td>
<td>General QOL</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>------------</td>
<td>----------------</td>
<td>----------</td>
<td>---------------</td>
<td>---------</td>
<td>------------</td>
<td>---------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Maxwell et al, 2013</td>
<td>Retrospective cohort</td>
<td>Screening program in the United Kingdom</td>
<td>First screening  A) Open biopsy (n=110) b) Needle sampling (n=1,374) C) No tissue sampling (n=2,703)</td>
<td>Re-attendance</td>
<td>Increased for C, but no change for A or B</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Repeat screening  A) Open biopsy (n=199) b) Needle sampling (n=1,052) C) No tissue sampling (n=4,009)</td>
<td></td>
<td>Decreased for A and B, but no change for C</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Tosteson et al, 2014</td>
<td>Nested case-control</td>
<td>Women participating in DMIST in the United States</td>
<td>A) FP (n=494) immediate B) FP 1-year after</td>
<td>STAI, EuroQOL EQ-5D</td>
<td>NR</td>
<td>Decreased A to B</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C) Normal (n=534) immediate D) Normal 1-year after</td>
<td></td>
<td>NR</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

*Both groups improved over time.

**Abbreviations:** 0=comparison studied but not statistically significantly different; CES-D=Center for Epidemiological Studies-Depression Scale; COS-BC=Consequences of Screening in Breast Cancer; DMIST=Digital Mammographic Imaging Screening Trial; FP=false-positive; HADS=Hospital Depression and Anxiety Scale; n=number; NEO-FFI=Neuroticism-Extraversion-Openness-Five Factor Inventory; NR = not reported; QOL=quality of life; SF-36=Short-form 36 Health Survey; STAI=Spielberger State-Trait Anxiety Inventory; TP=true positive; vs.=versus; WHOQOL=World Health Organization Quality of Life Assessment Instrument; WHQ=Women's Health Questionnaire.
Table 24. Models of Radiation Exposure, Breast Cancer Incidence, and Death

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Population; age, years</th>
<th>Method</th>
<th>Outcome measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hendrick, 2010</td>
<td>Modeling Study</td>
<td>U.S. based sources; 40-80</td>
<td>Theoretical estimates are based on long-term followup 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 exam.</td>
<td>Breast cancer cases and mortality</td>
<td>LAR of breast cancer incidence and mortality, per 100,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• 40 years: 5-7 cases; 1.3-1.7 deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• 50 years: 2-3 cases; 0.7-0.9 deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• 80 years: 0.1-0.2 cases; &lt;0.1 deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LARs of breast cancer incidence and mortality in women undergoing annual screening mammography, per 100,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Screening 40-80 years: 72-91 cases; 20-25 deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Screening 50-80 years: 31-40 cases; 10-12 deaths</td>
</tr>
<tr>
<td>Yaffe and Mainprize, 2011</td>
<td>Modeling study</td>
<td>U.S. based sources; 40-74</td>
<td>Model based on digital mammography and radiation exposure estimates of 3.7 mGy per exam</td>
<td>Estimated lifetime radiation induced breast cancer cases and deaths</td>
<td>Number of radiation induced breast cancer cases and deaths in 100,000 women</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Annual screen 40-49 years: 59 cases, 7.6 deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Annual 50-59 years: 27 cases, 3.1 deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Biennial 50-59 years: 14 cases, 1.6 deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Annual 40-59 years: 85 cases, 11 deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Annual 40-49 years, biennial to 59 years: 73 cases, 9 deaths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Annual 40-55 years, biennial to 74 years: 86 cases, 11 deaths</td>
</tr>
</tbody>
</table>

Abbreviations: LAR=lifetime attributable risk; mGy=milli Gray (unit of radiation); U.S.=United States.
Table 25. Systematic Reviews of Pain With Mammography

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Inclusion criteria</th>
<th>Databases; search dates</th>
<th>Number of studies (designs); number of participants</th>
<th>Methods for rating quality and synthesizing results</th>
<th>Results</th>
<th>Quality rating; limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whelehan et al, 2013</td>
<td>Studies of pain or discomfort of screening mammography and re-attendance.</td>
<td>MEDLINE, EMBASE, PsychINFO, CINAHL, ASSIA, Cochrane Database of Systematic Reviews, Sociological Abstracts, SSCI, SCI, and NHS online literature database; October 2012.</td>
<td>20 observational studies (most cross-sectional surveys); causation N=5,741, association N=NR.</td>
<td>Quality based on individual factors;* studies combined separately for causation vs. association; actual vs. intended re-attendance data were considered more valid.</td>
<td>Causation (7 studies) Response rates: 32-79% • Actual non-re-attendance indicating pain as the reason (5 studies): 11-46% • Intended future non-re-attendance due to pain (2 studies): 2.7% and 17.5% Association (15 studies) • Actual re-attendance (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-re-attenders vs. re-attenders in 2 of 3 studies (p=0.001 and p&lt;0.05). • Intended re-attendance (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.</td>
<td>Fair; unclear how study quality was used to formulate conclusions; did not describe characteristics of all included studies; did not assess publication bias.</td>
</tr>
<tr>
<td>2009 Review</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armstrong et al, 2007</td>
<td>Studies of risks of screening mammography for women in their 40s.</td>
<td>MEDLINE, Pre-MEDLINE, and the Cochrane Central Register of Controlled Trials; May 2005.</td>
<td>22 studies (3 RCTs, 5 prospective cohort, 1 retrospective cohort, 13 cross-sectional); N=13,008.</td>
<td>Centre for Evidence-based Medicine criteria; based on study design and rates of attrition; methods of synthesis NR.</td>
<td>• 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 pre-mammography anticipation of pain (4 studies).</td>
<td>Fair; no synthesis of data; unclear how study quality was used to formulate conclusions; study designs not pre-specified; did not assess publication bias.</td>
</tr>
</tbody>
</table>
## Table 25. Systematic Reviews of Pain With Mammography

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Inclusion criteria</th>
<th>Databases; search dates</th>
<th>Number of studies (designs); number of participants</th>
<th>Methods for rating quality and synthesizing results</th>
<th>Results</th>
<th>Quality rating; limitations</th>
</tr>
</thead>
</table>
| Miller et al, 2008<sup>10</sup> | 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 RCTs; N=1,771. | Quality (levels A, B, C) based on generation and concealment of allocation sequence, comparability of groups at baseline, intention-to-treat analysis, and double-blinding after allocation; heterogeneity of studies allowed qualitative synthesis only. | 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 one trial (mean VAS 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. one 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 20.34 vs. 34.94, p<0.0001). | Good; did not assess publication bias. |

*Factors include whether intended or actual re-attendance 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 re-attendance rate.

**Abbreviations:** ASSIA=Applied Social Sciences Index and Abstracts; CI=confidence interval; CINAHL=Cumulative Index to Nursing and Allied Health Literature; EMBASE=Excerpta Medica database; N=number; NHS=National Health Services; NR=not reported; OR=odds ratio; RCT=randomized, controlled trial; RR=relative risk; SCI=Science Citation Index; SSCI=Social Sciences Citation Index; VAS=Visual Analogue Scale; vs.=versus.
Table 26. Studies of Harms of Screening With Different Modalities

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Population; age, years; participants, n</th>
<th>Study years; comparison</th>
<th>Outcome measures</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haas et al, 2013&lt;sup&gt;14&lt;/sup&gt;</td>
<td>Case series</td>
<td>U.S., multisite hospital and outpatient centers; DM, 7,058; DM + T, 6,100</td>
<td>2011 to 2012; DM vs. DM + T</td>
<td>Recall rate (%); adjusted odds of recall</td>
<td>Recall, DM vs. DM + T by age (% relative change, 95% CI): Total 8.4 vs. 12; -29.7 (19.1 to 36.5), p&lt;.01; 40 to 49, 10.4 vs. 16.3; -35.8 (24.2 to 45.7), p&lt;.01; 50 to 59, 7.6 vs. 10.6; -28 (12.7 to 44.6); p&lt;.01; 60 to 69, 7.4 vs. 10.7; -30.3 (12.3 to 44.6), p=.01; ≥70, 6.7 vs 7.9; -15.4; NS</td>
<td>Adjusted recall OR, 0.62 (0.55 to 0.70); p&lt;.0001</td>
</tr>
<tr>
<td>Friedewald et al, 2014&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Post-intervention series</td>
<td>U.S., multicenter; mean age 57; DM, 281,187; DM + T, 173,663</td>
<td>2010 to 2012; DM vs. DM + T</td>
<td>Recall and biopsy rates, per 1,000</td>
<td>Recall, DM vs. DM + T (change, 95% CI): 107/1,000 vs. 91/1,000; -16.1 (-18.0 to -14.2), p&lt;0.001 Biopsy, DM vs. DM + T (change, 95% CI): 18.1/1,000 vs. 19.3/1,000; +1.3 (0.4 to 2.1), p=0.004</td>
<td></td>
</tr>
<tr>
<td>Rose et al, 2013&lt;sup&gt;66&lt;/sup&gt;</td>
<td>Case series</td>
<td>U.S., multisite community-based breast center; &gt;18 years; 18,202 DM and 10,878 DM + T</td>
<td>2011 to 2012; DM vs. DM + T</td>
<td>Recall rate, %</td>
<td>Recall, DM vs. DM + T by age (% relative change): &lt;50, 10.3% vs. 6.5% (-37.2); 50 to 64, 7.6% vs. 5.1% (-32.9); age &gt;64, 7.9% vs. 4.2% (-46.6); total, 8.7% vs. 5.5% (-37.5); NS</td>
<td></td>
</tr>
<tr>
<td>Ciatto et al, 2013&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Post-intervention series</td>
<td>Italy; population-based screening program (STORM); ≥48; 7,292</td>
<td>2011 to June 2012; biennial DM vs. DM + T</td>
<td>Recall rate, %</td>
<td>Recall, DM vs. DM + T: total, 141 (2%) vs. 73 (1%), p&lt;0.0001; age &lt;60, 89 (2.2%) vs. 41 (1%); age &gt;60, 52 (2%) vs. 32 (1%)</td>
<td></td>
</tr>
<tr>
<td>Skaane et al, 2013&lt;sup&gt;47&lt;/sup&gt;</td>
<td>Post-intervention series</td>
<td>Norway, Oslo screening program; 50 to 69; 12,631</td>
<td>2010 to 2011; DM vs. DM + T (biennial screening)</td>
<td>Recall rate, per 1,000</td>
<td>Recall, DM vs. DM + T: 61.1/1,000 vs. 53.1/1,000 (-13%); RR 0.85 (p&lt;0.001)</td>
<td></td>
</tr>
</tbody>
</table>

**Mammography ± clinical breast exam**

| Chiarelli et al, 2009<sup>15</sup> | Cohort | Canada; 40-69; 290,230 | 2002 to 2003; biennial M (n=57,715) vs. CBE + M (n=232,515) | Recall rate, % | Recall, M vs. CBE ± M: 6.5% vs. 8.7% (+2.2% for CBE), or 55/10,000 additional FP with CBE | Fair |

**Abbreviations:** CBE=clinical breast exam; CI=confidence interval; DM=digital mammography; FP=false-positive; M=mammography; n=sample size; NR=not reported; NS=not statistically significant; OR=odds ratio; RR=relative risk; STORM=Screening with Tomosynthesis or standard Mammography; T=tomosynthesis; U.S.=United States; vs.=versus.
Table 27. Examples of Specific Adverse Effects of Selected Treatments for Breast Cancer

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Adverse Effect and Rate*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surgery</strong></td>
<td></td>
</tr>
<tr>
<td>Mastectomy</td>
<td>Wound infection 3.8%; skin flap necrosis 10% to 18%; chronic chest wall pain &gt;10%. Other adverse effects include phantom breast syndrome, arm morbidity, seroma, pneumothorax, brachial plexopathy, lymphedema.</td>
</tr>
<tr>
<td>Lymph node biopsy</td>
<td>Average false-negative 8.4% (range: 0% to 29% across 69 studies); &gt;1% of patients experienced allergic reactions to dye used during the procedure in a trial of 5588 patients. 5% with sentinel node biopsy and 16% to 18% with axillary lymph node dissection following sentinel node biopsy develop clinical lymphedema.</td>
</tr>
</tbody>
</table>

| Radiation†          |                                                                                        |
|---------------------|                                                                                        |
| Dose: 50 Gy in 25 fractions over 5 weeks | Based on 1,854 women: among women with breast conserving surgery—breast shrinkage 25%; breast induration 18%; telangiectasia 5%; breast edema 10%; among women who received lymphatic radiotherapy, shoulder stiffness 9%; arm edema 12%. Adverse effects experienced by <5% of patients: symptomatic rib fracture, symptomatic lung fibrosis, ischemic heart disease, brachial plexopathy. |
| Dose: 41.6 Gy in 13 fractions over 5 weeks | Based on 750 women: among women with breast conserving surgery—breast shrinkage 27%; breast induration 24%; telangiectasia 6%; breast edema 11%; among women who received lymphatic radiotherapy—shoulder stiffness 11%; arm edema 17%. Adverse effects experienced by <5% of patients: symptomatic rib fracture, symptomatic lung fibrosis, ischemic heart disease, brachial plexopathy. |
| Dose: 40 Gy in 15 fractions over 3 weeks  | Based on 1,110 women: among women with breast conserving surgery—breast shrinkage 22%; breast induration 13%; breast edema 5%. Adverse effects experienced by <5% of patients: telangiectasia, symptomatic rib fracture, symptomatic lung fibrosis, ischemic heart disease, brachial plexopathy, shoulder stiffness and arm edema in women who received lymphatic radiotherapy. |
| Dose: 39 Gy in 13 fractions over 5 weeks  | Based on 737 women: among women with breast conserving surgery—breast shrinkage 23%; breast induration 18%; breast edema 7%; among women who received lymphatic radiotherapy—shoulder stiffness 9%; arm edema 7%. Adverse effects experienced by <5% of patients: telangiectasia, symptomatic rib fracture, symptomatic lung fibrosis, ischemic heart disease, brachial plexopathy. |

| Endocrine therapy   |                                                                                        |
|---------------------|                                                                                        |
| **Anastrozole**     | Anastrozole treatment for 5 years (% of patients with common adverse events in a trial of 3125 postmenopausal patients with localized invasive breast cancer): fatigue 19%; nausea and vomiting 13%; hot flushes 36%; mood disturbances 19%; musculoskeletal disorders 36%. Adverse effects experienced by <5% of patients: vaginal bleeding, vaginal discharge, ischemic cardiovascular disease, ischemic cerebrovascular events, venous thromboembolic events, deep venous thromboembolic events, carpal tunnel syndrome. |
| Goserelin 3.6 mg given subcutaneously every 28 days plus anastrozole 1 mg/day for a mean of 47.8 months (% of patients each adverse events in a trial of 453 premenopausal patients who had undergone primary surgery for stage I or II breast cancer): arthralgia 24.7%; bone pain 28.3%; fatigue 20.5%; depression, sleep disturbances 21.4%; nausea and vomiting 7.1%; morning stiffness 7.3%; hot flushes 5.5%. Adverse effects experienced by <5% of patients: fracture, cognitive disorder, dizziness, peripheral nerve disease, muscle cramp, fever, hypertonia, tachycardia, thrombosis, leg edema, cutaneous reaction skin disease, impaired vision, uterine polyp, periodontal disease. |

| **Letrozole**        | Letrozole 2.5 mg/day for five years (% of patients with adverse events in a trial of 3975 postmenopausal patients with hormone receptor positive breast cancer): hot flushes 33.5%; night sweats 13.9%; fracture 5.7%; arthralgia 20.3%; myalgia 6.4%. Adverse effects experienced by <5% of patients: cerebrovascular accident or transient ischemic attack, thromboembolic event, cardiac event, other cardiovascular events, vaginal bleeding. |
### Table 27. Examples of Specific Adverse Effects of Selected Treatments for Breast Cancer

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Adverse Effect and Rate*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tamoxifen</td>
<td></td>
</tr>
<tr>
<td>Tamoxifen 10 vs. 20 mg/day orally for 6 months&lt;sup&gt;236&lt;/sup&gt; (% of patients reporting adverse events which occurred in &gt;10% of patients in a trial of 30 women with breast cancer): hot flashes 30%; nausea 17%; pharyngitis 17%; fatigue 13%. Tamoxifen 20 mg/day orally for 5 years&lt;sup&gt;218&lt;/sup&gt; (% of patients with adverse reactions in a trial of 1422 patients with primary operable breast cancer over the course of 5 years): hot flashes 64%; vaginal discharge 30%; irregular menses 25%; fluid retention 32%; nausea 24%; skin changes 19%; diarrhea 11%; weight gain 38%; weight loss 22%. Adverse effects experienced by &lt;5% of patients: thromboembolic vein, death. Other serious adverse effects of tamoxifen include an increased risk of endometrial cancer and uterine sarcoma.</td>
<td></td>
</tr>
<tr>
<td>Exemestane</td>
<td></td>
</tr>
<tr>
<td>Exemestane 25 mg/day orally for 5 years&lt;sup&gt;237&lt;/sup&gt; (% of patients with adverse reactions in a trial of 4852 postmenopausal patients with hormone receptor positive breast cancer): flushes and sweats 35%; hypertension 6%; breast or nipple disorder 6%; vaginal dryness 7%; fractures 5%; joint disorders 36%; muscle disorders 11%; osteoporosis 10%; other musculoskeletal and connective disuse disorders 15%; headache 8%; dizziness 5%; other nervous system disorders 17%; depression 9%; sleep disorder or insomnia 13%; other psychiatric disorders 8%; hyperlipidemia 5%; weight increase 7%. Adverse effect experienced by &lt;5% of patients: arrhythmia, cardiac failure, myocardial ischemia or infarction, other cardiac disorders, embolism, peripheral arterial disease, venous thrombosis, other vascular disorders, endometrial abnormalities, genital or vaginal discharge, postmenopausal bleeding, vulvovaginal disorders, cerebrovascular insufficiency or infarction or thrombosis, nerve compression disorders, loss or reduction of libido, abnormal liver function tests, endocrine disorders, renal and urinary disorders.</td>
<td></td>
</tr>
<tr>
<td>Neoadjuvant/adjuvant chemotherapy</td>
<td></td>
</tr>
<tr>
<td>AC (doxorubicin/cyclophosphamide) followed by paclitaxel</td>
<td>Doxorubicin 60 mg/m&lt;sup&gt;2&lt;/sup&gt; by 5 to 15-minute IV infusion and cyclophosphamide 600 mg/m&lt;sup&gt;2&lt;/sup&gt; by 30 to 60-minute IV infusion every 3 weeks for 4 cycles followed by 175 mg/m&lt;sup&gt;2&lt;/sup&gt; paclitaxel by 1-hour IV infusion weekly for 12 doses&lt;sup&gt;238&lt;/sup&gt; (% of patients with common toxic effects and neuropathies resulting from the paclitaxel component of therapy in 1231 patients with lymph node-positive or high risk, lymph node-negative breast cancer after mastectomy or breast conserving surgery): grade 2, 3, or 4 neuropathy 27%. Adverse effects experienced by &lt;6% of patients: neutropenia, febrile neutropenia, infections, stomatitis, fatigue, myalgia, arthralgia, lacrimation.</td>
</tr>
<tr>
<td>TC (docetaxel and cyclophosphamide)</td>
<td>Docetaxel 75 mg/m&lt;sup&gt;2&lt;/sup&gt; and cyclophosphamide 600 mg/m&lt;sup&gt;2&lt;/sup&gt; as 1-hour IV infusion on day 1 of a 3-week cycle (4 cycles total)&lt;sup&gt;259&lt;/sup&gt; (frequency [%] of side effects in 506 patients with operable stage I to III invasive breast cancer after surgical excision of the primary tumor): anemia &lt;7%; neutropenia 63%; thrombocytopenia &lt;3%; asthenia &lt;79%; edema &lt;35%; fever 24%; infection &lt;20%; myalgia &lt;35%; nausea &lt;55%; phlebitis &lt;12%; stomatitis &lt;35%; vomiting &lt;16%.</td>
</tr>
</tbody>
</table>
Table 27. Examples of Specific Adverse Effects of Selected Treatments for Breast Cancer

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Adverse Effect and Rate*</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCH (docetaxel/carboplatin/trastuzumab) +/- pertuzumab</td>
<td>Docetaxel 75 mg/m² plus carboplatin administered at area under the plasma concentration curve x6 mg/mL/minute concurrently with trastuzumab at 2mg/kg every 3 weeks for 6 cycles followed by trastuzumab 6 mg/kg every 3 weeks to complete 1 year of treatment 260 (% of patients with adverse events in a trial of 1056 patients with HER2 positive early-stage breast cancer): irregular menses 26.5%; sensory neuropathy 36%; nail changes 28%; myalgia 38.9%; neutropenia 65.9%; leukopenia 48.2%; febrile neutropenia 9.6% neutropenic infection 11.2%; anemia 5.8% thrombocytopenia: 6.1%. Adverse effect experiences by &lt;5% of patients: arthralgia, fatigue, hand-foot syndrome, diarrhea, nausea, vomiting, motor neuropathy, renal failure, grade 3 or 4 creatinine elevation, leukemia. Trastuzumab at an initial dose of 8 mg/kg, followed by 6 mg/kg; pertuzumab at an initial dose of 840 mg, followed by 420 mg, carboplatin was administered at a dose of 6x area under the plasma concentration-time curve and docetaxel was given at 75 mg/m² 261 (% of patients with the most common adverse events during neoadjuvant treatment in a trial of 76 patients with HER2 positive breast cancer): diarrhea 72.4%; alopecia 53.9%; nausea 44.7%; neutropenia 48.7%; vomiting 39/5%; fatigue 42.1%; anemia 36.8%; mucosal inflammation 17.1%; constipation 15.8%; dyspepsia 22.4%; febrile neutropenia 17.1%; leukopenia 11.8%; anemia 17.1%; thrombocytopenia 11.8%. Adverse effects experienced by &lt;5% of patients: drug hypersensitivity, alanine aminotransferase increase. Results reported as % of patients with the most common adverse events during adjuvant treatment in a trial of 67 patients with HER2 positive breast cancer: radiation skin injury 10.4%; arthralgia 9%; hot flushes 6%; diarrhea 9%; fatigue 7.5%; musculoskeletal chest pain 7.5%; peripheral edema 6%; erythema 6%. Adverse effects experienced by &lt;5% of patients: headache, musculoskeletal pain, neutropenia.</td>
</tr>
</tbody>
</table>

Chemotherapy regimens (metastatic cancer)

| Paclitaxel (taxane) | Paclitaxel 80mg/m² weekly via 1 hour infusion until disease progression or limiting toxicity (HER-2 + patients also received trastuzumab 2mg/kg via 30 minute infusion following a 4 mg/kg loading dose administered over 90 minutes) 262 (% of patients with grade 3 or 4 nonhematologic toxicity in a trial of 577 patients metastatic breast cancer): infection 6%; diarrhea 5%; dyspepsia 7%; edema 6%; neurosensory 24%; neuromotor 9%; malaise/fatigue 6%; Hyperglycemia: 5%. Two treatment related deaths attributable to pneumonia, and one secondary malignancy also occurred all in patients without trastuzumab. Paclitaxel 175mg/m² every 3 weeks via 3 hour infusion until disease progression or limiting toxicity (HER-2 + patients also received trastuzumab 2mg/kg via 30 minute infusion following a 4 mg/kg loading dose administered over 90 minutes) 262 (% of patients with grade 3 or 4 nonhematologic toxicity in a trial of 158 patients metastatic breast cancer): neurosensory 12%; Hyperglycemia: 8%. Adverse effects experienced by <5% of patients: infection, diarrhea, dyspepsia, edema, neuromotor, malaise/fatigue. One secondary malignancy also occurred in a patient without trastuzumab. |
| Docetaxel (taxane) | Docetaxel 50 mg/m² as a 1-hour IV infusion on days 1 and 8 every 3 weeks (100 mg/m² per cycle) for a median of 5 cycles 263 (% of patients with adverse effects in a trial of 88 patients with metastatic breast cancer): neutropenia 94%; thrombocytopenia 11%; anemia 90%; alopecia 91%; asthenia 82%; skin 64%; diarrhea 62%; nausea 54%; vomiting 43%; stomatitis 54%; neurosensory 48%; infection 16%; weight gain 28%; myalgia 18%; hypersensitivity reactions 4%. |
| Capecitabine (anti metabolite) | Capecitabine 1250 mg/m² twice/day orally for 14 days followed by a 7-day rest period, continued for a maximum of 15 cycles 264 (% of patients with common adverse events in a trial of 126 patients with anthracycline and taxane pretreated metastatic breast cancer): hand-foot syndrome 71%; nausea 48%; asthenia 35%; vomiting 27%; neutropenia 26%; stomatitis 25%. |

*This is not a comprehensive list of all potential adverse effects and reflects only events where rates were available.
Table 27. Examples of Specific Adverse Effects of Selected Treatments for Breast Cancer

†Results reported as % of patients experiencing adverse event during 10 years followup in 1,854 women with completely excised invasive breast cancer after primary surgery followed by chemotherapy and endocrine treatment where prescribed.217

**Abbreviations:** Gy=gray (unit of radiation); HER2= human epidermal growth factor receptor 2; IV = intravenous; kg=kilograms; m=meter; mg=milligram.
Table 28. Summary of Evidence: Screening for Breast Cancer

<table>
<thead>
<tr>
<th>Key Question 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?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main findings from previous USPSTF reviews</strong></td>
</tr>
<tr>
<td><strong>Number/type of studies in update</strong></td>
</tr>
<tr>
<td>Screening reduced breast cancer mortality in RCTs for women age 39-49 (RR 0.85; 95% CI, 0.75 to 0.96; 8 trials); 50-59 (RR 0.86; 95% CI, 0.75 to 0.99; 6 trials); and 60-69 (0.68; 95% CI, 0.54 to 0.87; 2 trials); data were limited for 70-74.</td>
</tr>
<tr>
<td>Not included.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Key Question 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 factors, and screening interval?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main findings from previous USPSTF reviews</strong></td>
</tr>
<tr>
<td><strong>Number/type of studies in update</strong></td>
</tr>
<tr>
<td>Not included.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Table 28. Summary of Evidence: Screening for Breast Cancer

<table>
<thead>
<tr>
<th>Main findings from previous USPSTF reviews</th>
<th>Number/type of studies in update</th>
<th>Overall quality*</th>
<th>Limitations</th>
<th>Consistency</th>
<th>Applicability</th>
<th>Summary of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key Question 3. How does the effectiveness of routine breast cancer screening in reducing breast cancer–specific and all-cause mortality vary by different screening modality?</td>
<td>Not included.</td>
<td>No studies evaluated this question.</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>RCTs of mammography with or without CBE do not compare relative mortality reduction across the different modalities.</td>
</tr>
<tr>
<td>Key Question 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?</td>
<td>Not included.</td>
<td>2 case-series studies</td>
<td>Poor</td>
<td>No RCTs; comparability of groups not known.</td>
<td>Results are consistent.</td>
<td>Tumor size, stage, and node status did not differ between women screened with tomosynthesis + digital mammography compared with those receiving mammography alone in 2 case-series studies.</td>
</tr>
<tr>
<td>Key Question 5. What are the harms of routine mammography screening, and how do they differ by age, risk factor, and screening interval?</td>
<td>Analysis of BCSC data showed that younger women had more false-positives results; the cumulative risk for false-positive mammograms was 21% to 49% after 10 screens, and 56% for age 40-49; cumulative false-positive biopsy rate after 10 screens was 19%. Many women have anxiety and pain with mammography, but it is generally transient and not a deterrent. Estimates of overdiagnosis ranged from 0% to 50% in a review and 8 studies.</td>
<td>Analysis of BCSC data; 3 observational studies of cumulative false positive results; 4 systematic reviews and 10 studies of anxiety; 3 reviews of pain; 1 meta-analysis, 2 reviews, and 27 studies of overdiagnosis; 2 modeling studies of radiation exposure</td>
<td>Poor (radiation) to good (false-positive results)</td>
<td>Limitations vary by outcome; lack of studies for some outcomes (radiation); methodological diversity of studies (overdiagnosis); lack of RCTs; comparability of groups vary in observational studies.</td>
<td>Consistent in general</td>
<td>High clinical relevance.</td>
</tr>
</tbody>
</table>

- Younger women and those with risk factors had more false-positives results, and recommendations for additional imaging and biopsies.
- 10-year cumulative rates of false-positive mammography and biopsy results were higher for annual than biennial screening (mammography 61% vs. 41%; biopsy 7% vs. 5%); for women with heterogeneously dense or extremely dense breasts, women in their 40s, and those who used combination hormone therapy.
- Women with false-positive results were more distressed than women with negative results, and some women did not return for screening.
- Estimates of overdiagnosis based on trials ranged from 11% to 22%; estimates from other studies ranged from 0% to 54%.
- Deaths due to radiation induced cancer from screening with digital mammography was estimated through modeling as between 2 to 11 per 100,000 depending on age at onset and screening intervals.
### Table 28. Summary of Evidence: Screening for Breast Cancer

| Key Question 6. How do the harms of routine breast cancer screening vary by different screening modality? |
|---|---|---|---|---|---|
| Not included. | 6 observational studies | Poor | No RCTs; single studies; comparability of groups not known. | Lack of studies to access consistency | High clinical relevance | Tomosynthesis with mammography reduced recalls, but increased biopsies. |

**Abbreviations:** BCSC= Breast Cancer Surveillance Consortium; CBE=clinical breast exam; CI=confidence interval; CrI=credible interval; NA=not applicable; RCT=randomized controlled trial; RR=relative risk; U.S.=United States; USPSTF=United States Preventive Services Task Force.
Table 29. Summary of Results: Benefits and Harms of Mammography Screening

### Benefits of Mammography Screening

<table>
<thead>
<tr>
<th>Age, years</th>
<th>Reduction in breast cancer deaths from RCTs; RR (95% CI)</th>
<th>Breast cancer deaths prevented per 10,000 over 10 years (95% CI)</th>
<th>Reduction in breast cancer deaths from observational studies; RR (95% CI)</th>
<th>Reduction in all-cause deaths from RCTs; RR (95% CI)</th>
<th>Reduction in advanced breast cancer from RCTs; RR (95% CI)</th>
<th>Reduction in treatment morbidity from RCTs; RR (95% CI)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-49</td>
<td>0.88 (0.73 to 1.03)</td>
<td>4 (0 to 9)</td>
<td>0.74 (0.66 to 0.83); 0.56 (0.45 to 0.67)†</td>
<td>0.99 (0.94 to 1.06)</td>
<td>0.98 (0.74 to 1.37)</td>
<td>Screening results in more mastectomies 1.20 (1.11 to 1.30) and radiation 1.32 (1.16 to 1.50); the majority of cases from screening are DCIS and early stage.</td>
</tr>
<tr>
<td>50-59</td>
<td>0.86 (0.68 to 0.97)</td>
<td>5 to 8 (0 to 17)</td>
<td>0.86 (0.69 to 1.007)</td>
<td>1.02 (0.94 to 1.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>0.67 (0.54 to 0.83)</td>
<td>12 to 21 (3 to 32)</td>
<td>0.67 (0.55 to 0.91)</td>
<td>0.97 (0.90 to 1.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70-74</td>
<td>0.80 (0.51 to 1.28)</td>
<td>12 to 13 (0 to 32)</td>
<td>0.90 (0.46 to 1.78)</td>
<td>1.02 (0.94 to 1.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-69</td>
<td>0.78 (0.68 to 0.90)</td>
<td>6 to 13 (1 to 20)</td>
<td>0.78 (0.69 to 0.95)</td>
<td>0.78 (0.69 to 0.81)§</td>
<td>0.69 (0.57 to 0.83)║</td>
<td>0.62 (0.46 to 0.83)</td>
</tr>
</tbody>
</table>

*From meta-analyses of screening trials using two different methods of case accrual; long case accrual results are provided first, then short case accrual results.
†Based on trials of screening included in the meta-analysis.
‡Based on a study in Sweden, and a study in Canada (standardized mortality ratio), respectively.
§Based on seven incidence-based mortality studies.
║Based on eight case-control studies.

**Abbreviations:** CI=confidence interval; RR=relative risk.

### Harms of Mammography Screening

<table>
<thead>
<tr>
<th>Age, years</th>
<th>False-positive mammography*</th>
<th>Additional imaging recommended*</th>
<th>Biopsy recommended*</th>
<th>10-yr FP mammography rates (annual; biennial)</th>
<th>10-yr FP biopsy rates (annual; biennial)</th>
<th>Overdiagnosis estimates from RCTs (95% CI)?</th>
<th>Overdiagnosis estimates from screening programs†</th>
<th>Radiation exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-49</td>
<td>121.2</td>
<td>124.9</td>
<td>16.4</td>
<td>61%: 42%</td>
<td>7%: 5%</td>
<td>10.7 (9.3 to 12.2)</td>
<td>0 to 54% unadjusted 1 to 10% adjusted</td>
<td>Annual screening 40-55 years, biennial to 74 years: 86 cases, 11 deaths§</td>
</tr>
<tr>
<td>50-59</td>
<td>93.2</td>
<td>98.5</td>
<td>15.9</td>
<td>61%: 42%</td>
<td>9%: 6%</td>
<td>19.0 (15.2 to 22.7)</td>
<td>0 to 54% unadjusted 1 to 10% adjusted</td>
<td>Annual screening 40-55 years, biennial to 74 years: 86 cases, 11 deaths§</td>
</tr>
<tr>
<td>60-69</td>
<td>80.8</td>
<td>88.7</td>
<td>16.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70-74</td>
<td>69.6</td>
<td>79.0</td>
<td>17.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Number per 1,000 screened per screening round.
†From meta-analysis of screening trials using two different methods of case accrual; long case accrual results are provided first, then short case accrual results.
‡From EUROSCREEN review based on 13 studies overall and 6 studies adjusted for breast cancer risk and lead time.
§From a model of digital mammography.

**Abbreviations:** CI=confidence interval; FP=false positive.
Appendix A1. Search Strategies

Database: Ovid MEDLINE(R) without Revisions
Search Strategy:

1 exp mammography/
2 exp physical examination/
3 exp magnetic resonance imaging/
4 exp ultrasonography/
5 exp mass screening/
6 ((clinical or physical$ or manual$ or routin$ or (regular$ adj2 schedul$)) adj5 (breast$ adj2 exam$)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
7 2 or 3 or 4 or 5 or 6
8 exp breast/
9 exp breast diseases/di, ra, us, pa, ep, mo
10 8 or 9
11 7 and 10
12 1 or 11
13 exp Mortality/
14 exp death/
15 exp survival analysis/
16 exp survivors/
17 mo.fs.
18 exp life tables/
19 exp life expectancy/
20 13 or 14 or 15 or 16 or 17 or 18 or 19
21 12 and 20
22 ((III$ or IV$ or advanc$ or late) adj5 stag$).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
23 (tnm adj7 (t3 or t4 or n1 or n2 or n3 or n4 or n5 or n6 or m1)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
24 ((cancer$ or tumor$ or tumour$ or malig$ or adenocarcin$ or nepolas$) adj5 (advanc$ or spread$ or infiltrat$ or metasta$)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
25 22 or 23 or 24
26 12 and 25
27 exp "Outcome Assessment (Health Care)"/
Appendix A1. Search Strategies

28 12 and 27
29 exp incidence/
30 12 and 29
31 exp Neoplasm Metastasis/
32 exp neoplasm staging/
33 exp neoplasm grading/
34 31 or 32 or 33
35 12 and 34
36 26 or 28 or 30 or 35
37 exp Breast Neoplasms/
38 36 and 37
39 exp Mammography/ae, ct [Adverse Effects, Contraindications]
40 exp Physical Examination/ae, ct
41 exp Mass Screening/ae, ct [Adverse Effects, Contraindications]
42 40 or 41
43 10 and 42
44 13 or 43
45 exp Diagnostic Errors/
46 (overtest$ or overdiagnos$ or over-test$ or over-diagnos$).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
47 misdiagnos$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
48 (false$ adj (positiv$ or negativ$)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
49 ((incorrect$ or false$ or wrong$ or bias$ or mistake$ or error$ or erroneous$) adj3 (result$ or finding$ or outcome$ or test$ or diagnos$)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
50 ((inappropriat$ or unnecess$ or unneed$) adj3 (treat$ or surg$ or therap$ or procedur$ or biops$ or interven$ or regimen$)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
51 (observ$ adj3 bias$).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
52 iatrogen$.mp.
53 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52
Appendix A1. Search Strategies

54 12 and 53
55 exp "Wounds and Injuries"/ci, et [Chemically Induced, Etiology]
56 exp Stress, Psychological/
57 exp Prejudice/
58 exp Stereotyping/
59 55 or 56 or 57 or 58
60 12 and 59
61 44 or 54 or 60
62 exp "sensitivity and specificity"/
63 12 and 62
64 exp *Breast Neoplasms/di, pa, ra, us
65 63 and 64

Databases: EBM Reviews - Cochrane Central Register of Controlled Trials, Database of Abstracts Reviews of Effects, Health Technology Assessment, and NHS Economic Evaluation Database
Search Strategy:

1 (mammogra$ or magnetic resonance imag$ or mri or ultrasound$ or ultrasonog$ or screen$ or ((clinical or physical$ or manual$ or routin$ or (regular$ adj2 schedul$)) adj5 exam$)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
2 ((Breast$ or mammar$) adj5 (cancer$ or tumor$ or tumour$ or neoplas$ or carcino$ or adenocarcino$ or malig$ or metasta$)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
3 (Mortal$ or death$ or dead or dying or die or dies or died or surviv$ or life table$ or life expectanc$).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
4 1 and 2 and 3
5 ((III$ or IV$ or advanc$ or late) adj5 stag$).mp.
6 (tnm adj7 (t3 or t4 or n1 or n2 or n3 or n4 or n5 or n6 or m1)).mp.
7 ((cancer$ or tumor$ or tumour$ or malig$ or adenocarcin$ or nepolas$) adj5 (advanc$ or spread$ or infiltrat$ or metasta$)).mp.
8 ((cancer$ or Neoplas$ or tumor$ or tumour$ or malig$ or carcino$) adj5 (Metasta or staging$ or stage or stages or grading or grades or graded or grade)).mp.
9 ((outcome$ or ((treat$ or therap$) adj3 result$)) adj5 (evaluat$ or compar$ or assess$)).mp.
10 5 or 6 or 7 or 8
11 1 and 2 and 10
12 1 and 2 and 9
13 (overtest$ or overdiagnos$ or over-test$ or over-dianos$ or overtreat$ or over-treat$).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
14 misdiagnos$.mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
Appendix A1. Search Strategies

15 (false$ adj (positiv$ or negativ$)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
16 ((incorrect$ or false$ or wrong$ or bias$ or mistake$ or error$ or erroneous$) adj3 (result$ or finding$ or outcome$ or test$ or diagnos$)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
17 ((inappropriat$ or unnecess$ or unneed$) adj3 (treat$ or surg$ or therap$ or procedur$ or biops$ or interven$ or regimen$)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
18 (observ$ adj3 bias$).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
19 iatrogen$.mp.
20 (diagnos$ adj5 (erroneous$ or error$)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
21 ((anguish$ or (emotion$ or psych$ or mental$ or physical$ or social$ or socio$)) adj5 (stress$ or tension$ or pain$ or fear$ or undesir$ or unwant$)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
22 (harm$ or advers$).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
23 (prejudic$ or stereotyp$ or stigma$ or unfair$).mp.
24 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23
25 1 and 2 and 24
26 (((test$ or diagnos$ or screen$) adj3 (sensitiv$ and Specific)) or (fals$ adj3 (positiv$ or negativ$)) or ((type I or type II) adj5 error$) or (Predict$ or prognos$)) adj5 (Value$ or valid$ or accura$ or correct$)) or (ROC adj2 Curv$) or (Signal adj Noise adj3 Ratio$)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
27 1 and 2 and 26

Database: EBM Reviews - Cochrane Database of Systematic Reviews
Search Strategy:

1 (mammogra$ or magnetic resonance imag$ or mri or ultrasound$ or ultrasonog$ or screen$ or ((clinical or physical$ or manual$ or routin$ or (regular$ adj2 schedul$)) adj5 exam$)).mp. [mp=title, abstract, full text, keywords, caption text]
2 ((Breast$ or mammar$) adj5 (cancer$ or tumor$ or tumour$ or neoplas$ or carcino$ or adenocarcino$ or malig$ or metast$a$)).mp. [mp=title, abstract, full text, keywords, caption text]
3 (Mortal$ or death$ or dead or dying or die or dies or died or surviv$ or life table$ or life expectanc$).mp. [mp=title, abstract, full text, keywords, caption text]
4 1 and 2 and 3
5 ((Breast$ or mammar$) adj5 (cancer$ or tumor$ or tumour$ or neoplas$ or carcino$ or adenocarcino$ or malig$ or metast$a$) adj15 (Mortal$ or death$ or dead or dying or die or dies or
Appendix A1. Search Strategies

died or surviv$ or life table$ or life expectanc$)).mp. [mp=title, short title, abstract, full text, keywords, caption text]
6 1 and 5
7 ((mammogra$ or magnetic resonance imag$ or mri or ultrasound$ or ultrasonog$ or screen$ or ((clinical or physical$ or manual$ or routin$ or (regular$ adj2 schedul$)) adj5 exam$)) adj15 (Mortal$ or death$ or dead or dying or die or dies or died or surviv$ or life table$ or life expectanc$)).mp. [mp=title, short title, abstract, full text, keywords, caption text] (387)
8 2 and 7
9 6 or 8
10 ((mammogra$ or magnetic resonance imag$ or mri or ultrasound$ or ultrasonog$ or screen$ or ((clinical or physical$ or manual$ or routin$ or (regular$ adj2 schedul$)) adj5 exam$)) adj10 ((Breast$ or mammar$) adj5 (cancer$ or tumor$ or tumour$ or neoplas$ or carcino$ or adenocarcino$ or malig$ or metast$))).mp. [mp=title, short title, abstract, full text, keywords, caption text]
11 3 and 10
12 6 or 9 or 11
13 (((III$ or IV$ or advanc$ or late) adj5 stag$)).mp.
14 (tnm adj7 (t3 or t4 or n1 or n2 or n3 or n4 or n5 or n6 or m1)).mp.
15 ((cancer$ or tumor$ or tumour$ or malig$ or adenocarcin$ or nepolas$) adj5 (advanc$ or spread$ or infiltrat$ or metast$)).mp.
16 ((cancer$ or Neoplas$ or tumor$ or tumour$ or malig$ or carcino$) adj5 (Metasta or staging$ or stage or stages or grading or grades or graded or grade)).mp. (502)
17 ((outcome$ or ((treat$ or therap$) adj3 result$)) adj5 (evaluat$ or compar$ or assess$)).mp.
18 13 or 14 or 15 or 16
19 1 and 2 and 18
20 1 and 2 and 17
21 (overtest$ or overdiagnos$ or over-test$ or over-diagnos$ or overtreat$ or over-treat$).mp. [mp=title, abstract, full text, keywords, caption text]
22 misdiagnos$.mp. [mp=title, abstract, full text, keywords, caption text]
23 (false$ adj (positiv$ or negativ$)).mp. [mp=title, abstract, full text, keywords, caption text]
24 ((incorrect$ or false$ or wrong$ or bias$ or mistake$ or error$ or erroneous$) adj3 (result$ or finding$ or outcome$ or test$ or diagnos$)).mp. [mp=title, abstract, full text, keywords, caption text]
25 ((inappropriat$ or unnecess$ or unneed$) adj3 (treat$ or surg$ or therap$ or procedur$ or biops$ or interven$ or regimen$)).mp. [mp=title, abstract, full text, keywords, caption text]
26 (observ$ adj3 bias$).mp. [mp=title, abstract, full text, keywords, caption text]
27 iatrogen$.mp.
28 (diagnos$ adj5 (erroneous$ or error$)).mp. [mp=title, abstract, full text, keywords, caption text]
Appendix A1. Search Strategies

29 ((anguish$ or (emotion$ or psych$ or mental$ or physical$ or social$ or socio$)) adj5 (stress$ or tension$ or pain$ or fear$ or undesir$ or unwant$)).mp. [mp=title, abstract, full text, keywords, caption text]
30 (harm$ or advers$).mp. [mp=title, abstract, full text, keywords, caption text]
31 (prejudic$ or stereotyp$ or stigma$ or unfair$).mp.
32 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31
33 1 and 2 and 32
34 ((((test$ or diagnos$ or screen$) adj3 (sensitiv$ and Specific)) or (fals$ adj3 (positiv$ or negativ$)) or ((type I or type II) adj5 error$) or (Predict$ or prognos$)) adj5 (Value$ or valid$ or accura$ or correct$)) or (ROC adj2 Curv$) or (Signal adj Noise adj3 Ratio$)).mp. [mp=title, abstract, full text, keywords, caption text]
35 1 and 2 and 34
## Appendix A2. Inclusion and Exclusion Criteria

<table>
<thead>
<tr>
<th>Include</th>
<th>Exclude</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
<td>KQs 1–6: Women age ≥40 years.</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td>KQs 1, 2, 5: Screening mammography (all methods, i.e., film, digital, tomosynthesis). KQs 3, 4, 6: Screening mammography (all methods) combined with other modality; other screening modality (i.e., MRI, ultrasound).</td>
</tr>
<tr>
<td><strong>Comparisons</strong></td>
<td>KQs 1, 2, 5: Mammography in women ages 40–49 vs. 50–59 vs. 60–69 vs. 70–79 years (or other age comparisons); annual mammography vs. biennial vs. triennial vs. alternate intervals vs. none; presence of risk factor vs. none (e.g., family history, extremely dense breast tissue). KQs 3, 4, 6: Mammography (all types, i.e., film, digital, tomosynthesis) vs. other modality vs. mammography (all types) combined with other modality, including MRI and ultrasound; interval and age differences by modality.</td>
</tr>
<tr>
<td><strong>Outcomes: Benefits</strong></td>
<td>Final health outcomes: Reduced breast cancer mortality and all-cause mortality. Intermediate outcomes: Reduced incidence of advanced breast cancer and treatment-related morbidity (i.e., physical adverse effects of treatment, quality of life measures, and other measures of impairment).</td>
</tr>
<tr>
<td><strong>Outcomes: Harms</strong></td>
<td>False-positive findings; anxiety; adverse impact on quality of life; false-positive biopsy; false-negative findings; false reassurance; overdiagnosis; overtreatment; radiation exposure.</td>
</tr>
<tr>
<td><strong>Timing</strong></td>
<td>Immediate, short-term, and long-term outcomes; duration of followup.</td>
</tr>
<tr>
<td><strong>Setting</strong></td>
<td>Settings and populations of women applicable to U.S. primary care practice.</td>
</tr>
<tr>
<td><strong>Study Design</strong></td>
<td>Effectiveness: RCTs; prospective and retrospective cohort studies. Harms: RCTs, prospective and retrospective cohort studies, case-control studies, cross-sectional studies, systematic reviews, meta-analyses, and modeling studies; others considered.</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td>English-language abstracts (includes English-language abstracts of non English-language papers) and papers.</td>
</tr>
<tr>
<td><strong>Contextual Question 1</strong></td>
<td>U.S. rates of specific adverse effects of current treatment regimens for invasive breast cancer and DCIS from published sources and databases, obtained using a best evidence approach.</td>
</tr>
</tbody>
</table>
## Appendix A2. Inclusion and Exclusion Criteria

<table>
<thead>
<tr>
<th>Contextual Question</th>
<th>Include</th>
<th>Exclude</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Absolute incidence rates of DCIS and localized and advanced invasive breast cancer in screened and nonscreened populations in the United States from published sources and databases, obtained using a best evidence approach.</td>
<td>Non U.S. rates, older estimates (see search dates)</td>
</tr>
<tr>
<td>3</td>
<td>Descriptive papers of how women’s perceptions of the benefits and harms of breast cancer screening affect their clinical decision-making regarding breast cancer screening in the United States.</td>
<td>Studies of women in other countries; older studies (see search dates)</td>
</tr>
<tr>
<td>Data Sources</td>
<td>Ovid MEDLINE, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Breast Cancer Surveillance Consortium database.</td>
<td>Sources not listed as included</td>
</tr>
</tbody>
</table>
**Effectiveness Key Questions not included in the previous report (KQ 2 & 4):** RCTs and cohort studies published between 1996 and February 2014.  
**Harms (KQs 5, 6):** Studies published between 2008 and February 2014 and updates of earlier studies.  
**Contextual questions (1–3):** Studies published between 2010 and February 2014. | Studies published outside of the specified search dates that were not included in previous USPSTF systematic reviews. |

*Search dates vary because some key questions (KQs 1, 3, 5, 6) were included in the previous systematic review and require only an update of studies published since the previous search in 2008. Other key questions were not addressed by the previous review and require a search that covers a longer time period (KQ 2 & 4, and cohort studies for KQ 1 & 3). These searches extend to 1996 because this corresponds to the last time the USPSTF evaluated similar data and represents a period when practice was shifting to digital mammography in the United States. The contextual questions have a shorter time period for searches because they require current data.

**Abbreviations:** ADH=atypical ductal hyperplasia; ALH=atypical lobular hyperplasia; BRCA=breast cancer susceptibility gene; DCIS=ductal carcinoma in situ; Gy=gray (unit of absorbed radiation dose [1 Gy=100 rads]); KQ=key question; LCIS=lobular carcinoma in situ; MRI=magnetic resonance imaging; RCT=randomized, controlled trial; U.S.=United States; USPSTF=U.S. Preventive Services Task Force; vs.=versus.
Appendix A3. Literature Flow by Key Question

Abstracts of potentially relevant articles identified through MEDLINE and Cochrane* (N=12,004)

Excluded abstracts and background articles (n=9,970)

Full text articles reviewed for relevance to Key Questions (n=2,034)

Included studies=75†

Full text articles excluded=1,950
Wrong population=129
Wrong intervention=243
Wrong outcomes=532
Wrong study design for Key Question=214
Wrong publication type=307
Included in an included systematic review, not directly used=68
Wrong comparison=239
Review not meeting inclusion criteria =125
Studies outside of search dates=63
No original data to include; publication or dataset with longer followup, more complete data, or same data already included in review=30

Key Question 1.
3 updated RCTs
5 old RCTs, not updated
4 reviews (57 studies)
8 observational studies

Key Question 2.
4 RCTs
1 review (5 RCTs)
12 observational studies
BCSC analyses

Key Question 3.
No studies

Key Question 4.
2 observational studies

Key Question 5.
10 reviews (134 studies)
1 meta-analysis (3 RCTs)
40 observational studies
2 modeling studies
BCSC analyses

Key Question 6.
6 observational studies

*Cochrane databases include the Cochrane Central Register of Controlled Trials and the Cochrane Database of Systematic Reviews.
†Publications were used for multiple key questions; trials reported data in multiple publications; 84 publications were included.

Abbreviations: BCSC=Breast Cancer Surveillance Consortium; n=sample size; RCT=randomized controlled trial.
Appendix A4. List of Excluded Studies

3D mammography shows promise as next breast screening technique. Oncology. 2003;17(6):814, 900. **Exclusion:** wrong publication type.


Advances in breast imaging. Although mammography remains standard for breast cancer screening, several newer technologies are helping to fine-tune diagnosis. Harv Womens Health Watch. 2010;17(9):1-3. **Exclusion:** wrong publication type.


Computer-aided detection of malignancy with magnetic resonance imaging of the breast. Technol Eval Cent Asses Program Exec Summ. 2006;Executive Summary. 21(4):1-3. **Exclusion:** wrong publication type.

Computer-aided detection with full-field digital mammography. Technol Eval Cent Asses Program Exec Summ. 2006;Executive Summary. 21(3):1-3. **Exclusion:** review not meeting inclusion criteria.


Magnetic resonance imaging of the breast: differential diagnosis of a breast lesion to avoid biopsy. Tecnologica MAP Suppl. 2001:30-2. **Exclusion:** wrong outcomes.


Appendix A4. List of Excluded Studies

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: studies outside of search dates.

Exclusion: studies outside of search dates.

Exclusion: wrong study design for key question.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong outcomes.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

Appendix A4. List of Excluded Studies

**Exclusion**: wrong comparison.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong outcomes.


**Exclusion**: review not meeting inclusion criteria.


**Exclusion**: included in an included systematic review, not directly used.


**Exclusion**: wrong comparison.


**Exclusion**: wrong population.


**Exclusion**: included in an included systematic review, not directly used.


**Exclusion**: studies outside of search dates.


**Exclusion**: included in an included systematic review, not directly used.


**Exclusion**: wrong intervention.


**Exclusion**: included in an included systematic review, not directly used.


**Exclusion**: wrong comparison.


**Exclusion**: wrong publication type.


**Exclusion**: wrong comparison.


**Exclusion**: wrong intervention.


**Exclusion**: review not meeting inclusion criteria.


**Exclusion**: wrong study design for key question.

Autier P, Boniol M. The incidence of advanced breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. BMJ. 2011;343:d4411.

**Exclusion**: included in an included systematic review, not directly used.
## Appendix A4. List of Excluded Studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baines CJ, McFarlane DV, Miller AB. The role of the reference radiologist; estimates of inter-observer agreement and potential delay in cancer detection in the national breast screening study. Invest Radiol. 1990;25(9):971-6.</td>
<td>review not meeting inclusion criteria.</td>
</tr>
<tr>
<td>Baines CJ, Vidmar M, McKeown-Eyssen G, et al. Impact of menstrual phase on false-negative mammograms in the Canadian National Breast Screening Study. Cancer. 1997;80(4):720-4.</td>
<td>no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.</td>
</tr>
</tbody>
</table>
Appendix A4. List of Excluded Studies


Appendix A4. List of Excluded Studies

**Exclusion**: included in an included systematic review, not directly used.

**Exclusion**: wrong comparison.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong study design for key question.

**Exclusion**: wrong study design for key question.

**Exclusion**: review not meeting inclusion criteria.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong publication type.

**Exclusion**: wrong intervention.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong publication type.

**Exclusion**: review not meeting inclusion criteria.

**Exclusion**: wrong publication type.

**Exclusion**: wrong study design for key question.

**Exclusion**: wrong study design for key question.

BCA BCA. Health is Not Just Healthcare: Inequities in Breast Cancer. 2013
**Exclusion**: wrong publication type.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong study design for key question.
Appendix A4. List of Excluded Studies


**Exclusion:** study design for key question.


**Exclusion:** wrong outcomes.


**Exclusion:** studies outside of search dates.


**Exclusion:** wrong intervention.


**Exclusion:** outcomes.


**Exclusion:** study population.


**Exclusion:** included in an included systematic review, not directly used.


**Exclusion:** outcomes.


**Exclusion:** review not meeting inclusion criteria.


**Exclusion:** study design for key question.


**Exclusion:** comparison.


**Exclusion:** comparison.


**Exclusion:** outcomes.


**Exclusion:** outcomes.


**Exclusion:** publication type.


**Exclusion:** publication type.


**Exclusion:** publication type.


**Exclusion:** publication type.
Appendix A4. List of Excluded Studies


Berg WA, Mendelson EB. Technologist-performed handheld screening breast US imaging: how is it performed and what are the outcomes to date? Radiology. 2014;272(1):12-27. **Exclusion:** wrong publication type.


Bick U, Diekmann F. Digital mammography: what do we and what don't we know? Eur Radiol. 2007;17(8):1931-42. **Exclusion:** wrong publication type.


Appendix A4. List of Excluded Studies

**Exclusion:** wrong intervention.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** studies outside of search dates.

**Exclusion:** study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** wrong comparison.

**Exclusion:** wrong publication type.

**Exclusion:** wrong population.

**Exclusion:** wrong comparison.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong population.

**Exclusion:** wrong publication type.

**Exclusion:** wrong comparison.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** wrong study design for key question.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong population.

**Exclusion:** wrong population.

**Exclusion:** wrong publication type.

**Exclusion:** wrong comparison.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong comparison.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** included in an included systematic review, not directly used.
Appendix A4. List of Excluded Studies

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong intervention.

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.

Appendix A4. List of Excluded Studies

Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong outcomes.


Exclusion: wrong publication type.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong population.


Exclusion: wrong comparison.


Exclusion: wrong population.


Exclusion: wrong intervention.


Exclusion: wrong publication type.


Exclusion: wrong comparison.


Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong publication type.
Appendix A4. List of Excluded Studies

Exclusion: wrong publication type.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Carmichael AR. Obesity as a risk factor for development and poor prognosis of breast cancer. BJOG. 2006;113(10):1160-6.
Exclusion: review not meeting inclusion criteria.

Exclusion: studies outside of search dates.

Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong population.

Exclusion: wrong outcomes.
Appendix A4. List of Excluded Studies

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong population.

Exclusion: wrong intervention.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.
Appendix A4. List of Excluded Studies

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong comparison.

Exclusion: wrong comparison.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong intervention.

**Exclusion:** wrong publication type.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

Appendix A4. List of Excluded Studies

Exclusion: wrong comparison.


Exclusion: studies outside of search dates.


Exclusion: wrong outcomes.


Exclusion: wrong population.


Exclusion: wrong publication type.


Exclusion: wrong intervention.


Exclusion: wrong outcomes.


Exclusion: wrong intervention.


Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong publication type.


Exclusion: wrong population.


Exclusion: wrong study design for key question.


Exclusion: wrong study design for key question.


Exclusion: included in an included systematic review, not directly used.


Exclusion: studies outside of search dates.


Exclusion: included in an included systematic review, not directly used.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong outcomes.

**Exclusion:** studies outside of search dates.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** review not meeting inclusion criteria.

Consedine NS. A false-positive on screening mammography has a negative psychosocial impact up to 3 years after receiving the all clear. Evid Based Ment Health. 2013;16(4):115.
**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.
Appendix A4. List of Excluded Studies


**Exclusion:** wrong comparison.


**Exclusion:** studies outside of search dates.


**Exclusion:** wrong comparison.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong comparison.


**Exclusion:** wrong study design for key question.


**Exclusion:** review not meeting inclusion criteria.


**Exclusion:** wrong comparison.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong publication type.


**Exclusion:** wrong comparison.


**Exclusion:** wrong intervention.


**Exclusion:** wrong population.


**Exclusion:** wrong study design for key question.
Appendix A4. List of Excluded Studies

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Das B, Feuer EJ, Mariotto A. Geographic association between mammography use and mortality reduction in the US. Cancer Causes Control. 2005;16(6):691-9.
Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: studies outside of search dates.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.
Appendix A4. List of Excluded Studies

Exclusion: studies outside of search dates.

Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

Exclusion: review not meeting inclusion criteria.

de Koning HJ. Why improvement in survival of screen-detected cases is not necessarily equivalent to benefit? Breast. 2003;12(5):299-301.  
Exclusion: wrong publication type.

Exclusion: wrong comparison.

Exclusion: wrong intervention.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong comparison.

Exclusion: wrong comparison.

Exclusion: wrong population.

Exclusion: wrong population.

Exclusion: wrong population.

Exclusion: wrong publication type.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong comparison.

Exclusion: wrong population.

Exclusion: wrong population.
Appendix A4. List of Excluded Studies


**Exclusion**: wrong outcomes.


**Exclusion**: wrong comparison.


**Exclusion**: wrong intervention.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong publication type.


**Exclusion**: wrong intervention.


**Exclusion**: included in an included systematic review, not directly used.


**Exclusion**: wrong population.


**Exclusion**: wrong intervention.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong population.


**Exclusion**: wrong publication type.


**Exclusion**: wrong intervention.


**Exclusion**: wrong publication type.

Dissemination CfRa. Positron emission tomography: systematic review. PET as a diagnostic test in breast cancer (Structured abstract). Database of Abstracts of Reviews of Effects. 2014(1)

**Exclusion**: wrong publication type.


**Exclusion**: wrong population.
Appendix A4. List of Excluded Studies

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong population.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: studies outside of search dates.

Exclusion: included in an included systematic review, not directly used.

Exclusion: included in an included systematic review, not directly used.
Appendix A4. List of Excluded Studies


**Exclusion:** included in an included systematic review, not directly used.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong intervention.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong publication type.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong publication type.


**Exclusion:** wrong publication type.


**Exclusion:** wrong outcomes.


**Exclusion:** study design for key question.


**Exclusion:** intervention.


**Exclusion:** wrong outcomes.


**Exclusion:** study design for key question.


**Exclusion:** study design for key question.


**Exclusion:** not meeting inclusion criteria.


**Exclusion:** wrong publication type.

Appendix A4. List of Excluded Studies

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: wrong study design for key question.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong comparison.

Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong outcomes.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong publication type.

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong outcomes.
Appendix A4. List of Excluded Studies

Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong intervention.


Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong publication type.


Exclusion: studies outside of search dates.


Exclusion: wrong publication type.


Exclusion: wrong publication type.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong population.


Exclusion: wrong outcomes.


Exclusion: wrong study design for key question.


Exclusion: wrong publication type.


Exclusion: wrong study design for key question.


Exclusion: wrong publication type.

Appendix A4. List of Excluded Studies

Exclusion: wrong study design for key question.


Exclusion: wrong study design for key question.


Exclusion: wrong comparison.


Exclusion: wrong study design for key question.


Exclusion: wrong publication type.


Exclusion: wrong intervention.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong population.


Exclusion: wrong publication type.


Exclusion: wrong publication type.


Exclusion: wrong population.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong study design for key question.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong comparison.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong publication type.

**Exclusion:** no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong publication type.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.
Appendix A4. List of Excluded Studies


Appendix A4. List of Excluded Studies

**Exclusion**: wrong study design for key question.


**Exclusion**: wrong publication type.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong intervention.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong intervention.


**Exclusion**: studies outside of search dates.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong intervention.


**Exclusion**: review not meeting inclusion criteria.


**Exclusion**: review not meeting inclusion criteria.


**Exclusion**: review not meeting inclusion criteria.


**Exclusion**: wrong population.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong intervention.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong outcomes.
Appendix A4. List of Excluded Studies


Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. Radiology. 2001;220(3):781-6. **Exclusion:** study design for key question. **Exclusion:** wrong study design for key question.


Frisell J, Eklund G, Hellström L, et al. The Stockholm breast cancer screening trial - 5-year results and stage at discovery. Breast Cancer Res Treat. 1989;13(1):79-87. **Exclusion:** no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.


Appendix A4. List of Excluded Studies


Garnick MB. Screening and detection of breast cancer and prostate cancer. JAMA. 2010;303(11):1033; author reply -4. **Exclusion:** wrong publication type.
Appendix A4. List of Excluded Studies

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: studies outside of search dates.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong population.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong publication type.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.
Appendix A4. List of Excluded Studies

Exclusion: wrong outcomes.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong comparison.

Exclusion: wrong comparison.

Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

Exclusion: wrong publication type.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong publication type.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Appendix A4. List of Excluded Studies

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong comparison.

Exclusion: studies outside of search dates.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong population.

Exclusion: wrong outcomes.
Appendix A4. List of Excluded Studies


Habermann EB, Virmig BA, Riley GF, et al. The impact of a change in Medicare reimbursement policy and HEDIS measures on stage at diagnosis among Medicare HMO and fee-for-service female breast cancer patients. Med Care. 2007;45(8):761-6. **Exclusion:** wrong comparison.


Appendix A4. List of Excluded Studies


Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong intervention.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong publication type.

Hall FM. Mammographically determined breast density and cancer risk. Radiology. 2008;248(3):1083; author reply

Exclusion: wrong publication type.


Exclusion: wrong publication type.


Exclusion: wrong outcomes.


Exclusion: wrong publication type.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong publication type.


Exclusion: wrong population.


Exclusion: wrong study design for key question.


Exclusion: wrong outcomes.


Exclusion: wrong study design for key question.


Exclusion: wrong population.


Exclusion: wrong study design for key question.
Appendix A4. List of Excluded Studies

Exclusion: wrong intervention.


Exclusion: wrong publication type.


Exclusion: wrong publication type.


Exclusion: wrong publication type.


Exclusion: wrong comparison.


Exclusion: wrong publication type.


Exclusion: wrong population.


Exclusion: wrong publication type.


Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong comparison.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.

Helmo BM, Lindberg L, Brodersen J. Alternative estimates for the likelihood that a woman with screen-detected breast cancer has had her "life saved" by that screening. Arch Intern Med. 2012;172(8):672.

Exclusion: wrong publication type.


Exclusion: wrong publication type.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.
Appendix A4. List of Excluded Studies

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong population.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong population.

Exclusion: wrong study design for key question.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong outcomes.
Appendix A4. List of Excluded Studies

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong comparison.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong comparison.

Exclusion: review not meeting inclusion criteria.

Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

Appendix A4. List of Excluded Studies

Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: studies outside of search dates.


Exclusion: review not meeting inclusion criteria.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong publication type.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong publication type.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong publication type.


Exclusion: wrong intervention.


Exclusion: wrong publication type.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong outcomes.


Exclusion: wrong study design for key question.
Appendix A4. List of Excluded Studies

**Exclusion**: wrong study design for key question.

**Exclusion**: wrong intervention.

**Exclusion**: wrong intervention.

**Exclusion**: wrong comparison.

**Exclusion**: wrong intervention.

**Exclusion**: wrong intervention.

**Exclusion**: wrong population.

**Exclusion**: wrong study design for key question.

**Exclusion**: wrong comparison.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong population.

Institute for Clinical Systems I. Computer aided detection of breast cancer (Structured abstract). Health Technology Assessment Database. 2014(1)
**Exclusion**: wrong publication type.

Institute for Clinical Systems I. Magnetic resonance imaging (MRI) for the detection of breast abnormalities (Structured abstract). Health Technology Assessment Database. 2014(1)
**Exclusion**: wrong publication type.

**Exclusion**: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

**Exclusion**: review not meeting inclusion criteria.

**Exclusion**: wrong outcomes.
Appendix A4. List of Excluded Studies


Exclusion: wrong outcomes.


Exclusion: wrong publication type.


Exclusion: wrong publication type.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong intervention.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong comparison.


Exclusion: wrong study design for key question.


Exclusion: wrong intervention.


Exclusion: wrong publication type.


Exclusion: wrong intervention.


Exclusion: studies outside of search dates.


Exclusion: wrong intervention.


Exclusion: wrong comparison.


Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.
Appendix A4. List of Excluded Studies

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong population.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong intervention.

Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong population.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong population.

Exclusion: wrong comparison.

Exclusion: included in an included systematic review, not directly used.
Appendix A4. List of Excluded Studies


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: included in an included systematic review, not directly used.


Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong study design for key question.


Exclusion: wrong study design for key question.


Exclusion: wrong publication type.

Jørgensen KJ. Mammography screening is not as good as we hoped. Maturitas. 2010;65(1):1-2.

Exclusion: wrong publication type.


Exclusion: review not meeting inclusion criteria.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong publication type.

Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong publication type.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.
Appendix A4. List of Excluded Studies


**Exclusion**: wrong population.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong intervention.


**Exclusion**: wrong publication type.


**Exclusion**: wrong population.


**Exclusion**: wrong comparison.


**Exclusion**: included in an included systematic review, not directly used.


**Exclusion**: wrong publication type.


**Exclusion**: wrong comparison.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong publication type.


**Exclusion**: wrong intervention.


**Exclusion**: wrong publication type.


**Exclusion**: wrong comparison.


**Exclusion**: studies outside of search dates.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong population.
Appendix A4. List of Excluded Studies


Keen JD. Promoting screening mammography: insight or uptake? J Am Board Fam Med. 2010;23(6):775-82. **Exclusion:** wrong publication type.


Keen JD, Keen JE. How does age affect baseline screening mammography performance measures? A decision model. BMC Med Inform Decis Mak. 2008;8:40. **Exclusion:** wrong study design for key question.

Keen JD, Keen JE. What is the point: will screening mammography save my life? BMC Med Inform Decis Mak. 2009;9:18. **Exclusion:** wrong study design for key question.


Kerlikowske K. Efficacy of screening mammography among women aged 40 to 49 years and 50 to 69 years: comparison of relative and absolute benefit. J Natl Cancer Inst. 1997;Monographs.(22):79-86. **Exclusion:** review not meeting inclusion criteria.


Appendix A4. List of Excluded Studies


Appendix A4. List of Excluded Studies


Kmietowicz Z. Mammography reduces risk of breast cancer death by 28%, study finds. BMJ. 2014;348:g4062. **Exclusion**: wrong publication type.


Appendix A4. List of Excluded Studies


Kopans DB. Screening mammography for women age 40 to 49 years. Ann Intern Med. 2007;147(10):740-1; author reply 1. **Exclusion:** wrong publication type.


Appendix A4. List of Excluded Studies

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong population.

**Exclusion:** wrong population.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

Appendix A4. List of Excluded Studies

**Exclusion**: wrong comparison.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong population.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong publication type.

Lambertz CK. Pre-medication with Acetaminophen for Screening Mammography: Gonzaga University; 1998. Book

**Exclusion**: included in an included systematic review, not directly used.


**Exclusion**: studies outside of search dates.


**Exclusion**: studies outside of search dates.


**Exclusion**: included in an included systematic review, not directly used.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong publication type.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong intervention.
Appendix A4. List of Excluded Studies

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong population.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong population.

Exclusion: wrong outcomes.

Exclusion: review not meeting inclusion criteria.

Exclusion: review not meeting inclusion criteria.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong outcomes.

Exclusion: wrong population.
**Exclusion:** wrong comparison.

**Exclusion:** wrong intervention.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** studies outside of search dates.

**Exclusion:** studies outside of search dates.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** studies outside of search dates.

**Exclusion:** wrong publication type.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.
Appendix A4. List of Excluded Studies


Exclusion: wrong intervention.


Exclusion: wrong comparison.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong intervention.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong intervention.


Exclusion: wrong comparison.


Exclusion: wrong publication type.


Exclusion: wrong outcomes.


Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.


Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.


Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong publication type.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** studies outside of search dates.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong comparison.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong comparison.

**Exclusion:** wrong population.

**Exclusion:** wrong comparison.

**Exclusion:** included in an included systematic review, not directly used.
Appendix A4. List of Excluded Studies


Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong outcomes.


Exclusion: wrong study design for key question.


Exclusion: wrong population.


Exclusion: wrong comparison.


Exclusion: wrong publication type.


Exclusion: wrong outcomes.


Exclusion: studies outside of search dates.


Exclusion: studies outside of search dates.


Exclusion: wrong publication type.


Exclusion: wrong intervention.


Exclusion: wrong comparison.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong study design for key question.


Exclusion: wrong comparison.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong comparison.


Exclusion: wrong publication type.
Appendix A4. List of Excluded Studies


Marchick J, Henson DE. Correlations between access to mammography and breast cancer stage at diagnosis. Cancer. 2005;103(8):1571-80. **Exclusion:** wrong outcomes.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong outcomes.

Marquez Cruz MD, Marquez Calderon S. Diagnostic performance of digital mammography in breast cancer screening (Structured abstract). Health Technology Assessment Database. 2014(1)
**Exclusion:** wrong comparison.

**Exclusion:** wrong intervention.

**Exclusion:** studies outside of search dates.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong population.

Mayor S. Five year survival rate of women whose cancer is detected by screening has risen to 96.4%. BMJ. 2008;336(7658):1398-9.
**Exclusion:** wrong publication type.
Appendix A4. List of Excluded Studies

Mayor S. Mammography screening has little or no effect on breast cancer deaths, Swedish data indicate. BMJ. 2012;345:e4847.  
**Exclusion:** wrong publication type.

Mayor S. Number of breast cancer cases detected by screening has doubled in a decade in England. BMJ. 2009;338:b414.  
**Exclusion:** wrong outcomes.

**Exclusion:** wrong population.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong population.
Appendix A4. List of Excluded Studies


Exclusion: wrong comparison.


Exclusion: wrong intervention.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong comparison.


Exclusion: wrong publication type.


Exclusion: wrong publication type.


Exclusion: wrong population.

Medical Advisory S. Screening mammography for women aged 40 to 49 years at average risk for breast cancer: an evidence-based analysis (Structured abstract). Health Technology Assessment Database. 2014(1)

Exclusion: review not meeting inclusion criteria.


Exclusion: wrong intervention.


Exclusion: wrong outcomes.


Exclusion: wrong publication type.


Exclusion: wrong publication type.


Exclusion: included in an included systematic review, not directly used.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.

Appendix A4. List of Excluded Studies

Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: studies outside of search dates.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

Exclusion: wrong publication type.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** studies outside of search dates.

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** studies outside of search dates.

Miller D, Livingstone V, Herbison PG. Interventions for relieving the pain and discomfort of screening mammography. Cochrane Database Syst Rev. 2009(1)  
**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong publication type.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong population.

**Exclusion:** wrong intervention.

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong study design for key question.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.
Appendix A4. List of Excluded Studies

Morland B, Lund Haheim L, Linnestad K. Mammography screening (Structured abstract). Health Technology Assessment Database. 2014(1)
Exclusion: review not meeting inclusion criteria.

Exclusion: wrong study design for key question.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: studies outside of search dates.

Exclusion: review not meeting inclusion criteria.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong study design for key question.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Moy L, Noz ME, Maguire GQ, Jr., et al. Role of fusion of prone FDG-PET and magnetic resonance imaging of the breasts in the evaluation of breast cancer. Breast J. 2010;16(4):369-76.
Exclusion: wrong outcomes.
**Exclusion:** wrong intervention.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

Mundy L, Braunack-Mayer A, Merlin T. The role of scintimammography in the diagnosis of breast cancer (Structured abstract). Health Technology Assessment Database. 2014(1)  
**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

Murdy V, Pears R, Ball J, et al. An audit of screening for familial breast cancer before 50 years in the South Thames Region - have we got it right? Fam Cancer. 2004;3(1):29-34.  
**Exclusion:** wrong study design for key question.

**Exclusion:** wrong study design for key question.

**Exclusion:** no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.
Appendix A4. List of Excluded Studies


Appendix A4. List of Excluded Studies

**Exclusion**: wrong population.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong comparison.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong population.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong intervention.


**Exclusion**: wrong intervention.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong comparison.


**Exclusion**: review not meeting inclusion criteria.


**Exclusion**: wrong comparison.


**Exclusion**: wrong comparison.


**Exclusion**: wrong comparison.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong comparison.


**Exclusion**: wrong publication type.
Appendix A4. List of Excluded Studies


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong publication type.


**Exclusion:** wrong intervention.


**Exclusion:** wrong publication type.


**Exclusion:** wrong outcomes.


**Exclusion:** review not meeting inclusion criteria.


**Exclusion:** wrong intervention.


**Exclusion:** wrong intervention.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong comparison.


**Exclusion:** wrong outcomes.


**Exclusion:** studies outside of search dates.


**Exclusion:** studies outside of search dates.


**Exclusion:** wrong publication type.

Appendix A4. List of Excluded Studies

**Exclusion:** no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.


**Exclusion:** wrong publication type.


**Exclusion:** studies outside of search dates.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong outcomes.


**Exclusion:** studies outside of search dates.


**Exclusion:** wrong publication type.


**Exclusion:** wrong comparison.


**Exclusion:** wrong study design for key question.


**Exclusion:** included in an included systematic review, not directly used.


**Exclusion:** wrong publication type.


**Exclusion:** wrong study design for key question.


**Exclusion:** included in an included systematic review, not directly used.


**Exclusion:** wrong publication type.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong population.


**Exclusion:** wrong comparison.


**Exclusion:** studies outside of search dates.


**Exclusion:** studies outside of search dates.


**Exclusion:** wrong comparison.


**Exclusion:** wrong population.


**Exclusion:** studies outside of search dates.


**Exclusion:** wrong comparison.
Appendix A4. List of Excluded Studies

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong comparison.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong intervention.

Exclusion: wrong comparison.

Exclusion: wrong comparison.
O'Regan RM. Do tumors detected by mammography screening have a favorable prognosis? JAMA. 2004;292(9):1062-3.

Exclusion: wrong publication type.

Exclusion: review not meeting inclusion criteria.

Exclusion: review not meeting inclusion criteria.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong population.

Exclusion: wrong population.
Appendix A4. List of Excluded Studies

Exclusion: wrong outcomes.

Exclusion: included in an included systematic review, not directly used.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong population.

Exclusion: included in an included systematic review, not directly used.

Exclusion: included in an included systematic review, not directly used.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong comparison.

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: wrong study design for key question.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong publication type.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong study design for key question.
Appendix A4. List of Excluded Studies

Exclusion: studies outside of search dates.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong outcomes.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong population.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong publication type.

Exclusion: wrong study design for key question.
Appendix A4. List of Excluded Studies

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** wrong intervention.

**Exclusion:** wrong comparison.

**Exclusion:** wrong intervention.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong study design for key question.

**Exclusion:** no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** wrong intervention.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.
Appendix A4. List of Excluded Studies


Appendix A4. List of Excluded Studies

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong publication type.

Exclusion: studies outside of search dates.

Exclusion: wrong comparison.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong intervention.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.
Appendix A4. List of Excluded Studies


Poulos A, Rickard M. Compression in mammography and the perception of discomfort. Australas Radiol. 1997;41(3):247-52. **Exclusion:** included in an included systematic review, not directly used.


Printz C. Advanced cancer risk similar with biennial vs. annual mammography in women aged 50 to 74. Cancer. 2013;119(16):2949. **Exclusion:** wrong publication type.


Appendix A4. List of Excluded Studies

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong population.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong comparison.

**Exclusion:** review not meeting inclusion criteria.
Appendix A4. List of Excluded Studies


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong population.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong population.


Exclusion: wrong intervention.


Exclusion: wrong publication type.


Exclusion: wrong outcomes.


Exclusion: wrong intervention.


Exclusion: wrong outcomes.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong population.
Appendix A4. List of Excluded Studies


Appendix A4. List of Excluded Studies

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong population.

**Exclusion:** wrong publication type.

**Exclusion:** wrong comparison.

**Exclusion:** no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong population.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** wrong population.

**Exclusion:** wrong intervention.

**Exclusion:** review not meeting inclusion criteria.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** wrong comparison.
Appendix A4. List of Excluded Studies


Appendix A4. List of Excluded Studies

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** included in an included systematic review, not directly used.

**Exclusion:** wrong comparison.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** studies outside of search dates.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.
Appendix A4. List of Excluded Studies

Exclusion: wrong intervention.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong study design for key question.

Exclusion: wrong population.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong intervention.

Appendix A4. List of Excluded Studies

**Exclusion**: wrong study design for key question.


**Exclusion**: wrong outcomes.

Scotland NHSQI. Comparison of digital mammography and film screen mammography. 2004

**Exclusion**: wrong outcomes.


**Exclusion**: wrong population.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong comparison.


**Exclusion**: wrong comparison.


**Exclusion**: studies outside of search dates.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong population.


**Exclusion**: wrong population.

Semiglazov VF, Manikhas AG, Moiseenko VM, et al. [Results of a prospective randomized investigation [Russia (St.Petersburg)/WHO] to evaluate the significance of self-examination for the early detection of breast cancer]. Vopr Onkol. 2003;49(4):434-41.

**Exclusion**: wrong outcomes.


**Exclusion**: review not meeting inclusion criteria.


**Exclusion**: wrong publication type.


**Exclusion**: wrong outcomes.
Appendix A4. List of Excluded Studies

Exclusion: wrong study design for key question.

Seppanen J, Heinavaara S, Hakulinen T. Influence of alternative mammographic screening scenarios on breast cancer incidence predictions (Finland). Cancer Causes Control. 2006;17(9):1135-44.
Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong population.

Exclusion: studies outside of search dates.

Exclusion: studies outside of search dates.

Exclusion: wrong comparison.

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: review not meeting inclusion criteria.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong publication type.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: included in an included systematic review, not directly used.
Appendix A4. List of Excluded Studies


Skaane P. Nationwide mammography screening registers: the best way to go for mortality evaluation - but there is no free lunch. Acta Radiol. 2011;52(3):235. **Exclusion:** wrong publication type.


Skaane P. Ultrasonography as adjunct to mammography in the evaluation of breast tumors. Acta Radiol Suppl. 1999;420:1-47. **Exclusion:** wrong comparison.

Appendix A4. List of Excluded Studies


Appendix A4. List of Excluded Studies


Sterns EE. Relation between clinical and mammographic diagnosis of breast problems and the cancer/biopsy rate. Can J Surg. 1996;39(2):128-32. **Exclusion**: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.
Appendix A4. List of Excluded Studies

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: wrong study design for key question.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong population.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: wrong study design for key question.

Exclusion: included in an included systematic review, not directly used.

Exclusion: wrong comparison.

Exclusion: wrong population.
Appendix A4. List of Excluded Studies

**Exclusion**: wrong publication type.

**Exclusion**: studies outside of search dates.

**Exclusion**: wrong publication type.

**Exclusion**: studies outside of search dates.

**Exclusion**: studies outside of search dates.

**Exclusion**: studies outside of search dates.

**Exclusion**: studies outside of search dates.

**Exclusion**: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.

**Exclusion**: wrong comparison.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong population.

**Exclusion**: wrong study design for key question.

**Exclusion**: wrong intervention.

**Exclusion**: wrong outcomes.

**Exclusion**: wrong study design for key question.

Appendix A4. List of Excluded Studies

**Exclusion**: wrong outcomes.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong comparison.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong intervention.


**Exclusion**: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong intervention.


**Exclusion**: wrong intervention.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong comparison.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong intervention.


**Exclusion**: wrong intervention.


**Exclusion**: wrong comparison.
Appendix A4. List of Excluded Studies

Exclusion: wrong outcomes.

Exclusion: wrong population.

Exclusion: wrong publication type.

Exclusion: wrong comparison.

Exclusion: wrong publication type.

Exclusion: wrong intervention.

Exclusion: wrong comparison.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong population.

Exclusion: wrong population.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong intervention.
Appendix A4. List of Excluded Studies


Exclusion: review not meeting inclusion criteria.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong population.


Exclusion: wrong comparison.


Exclusion: wrong publication type.


Exclusion: wrong outcomes.


Exclusion: wrong intervention.


Exclusion: wrong outcomes.


Exclusion: wrong intervention.


Exclusion: wrong intervention.


Exclusion: wrong comparison.


Exclusion: wrong population.


Exclusion: wrong intervention.


Exclusion: wrong comparison.


Exclusion: wrong comparison.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.
Appendix A4. List of Excluded Studies


Exclusion: studies outside of search dates.


Exclusion: wrong intervention.


Exclusion: wrong outcomes.


Exclusion: wrong publication type.


Exclusion: wrong outcomes.


Exclusion: wrong comparison.


Exclusion: wrong comparison.


Exclusion: wrong intervention.


Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.


Exclusion: wrong outcomes.


Exclusion: wrong study design for key question.


Exclusion: wrong outcomes.


Exclusion: wrong study design for key question.


Exclusion: no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.


Exclusion: review not meeting inclusion criteria.


Exclusion: wrong comparison.


Exclusion: wrong comparison.
Appendix A4. List of Excluded Studies

Exclusion: wrong comparison.

Exclusion: wrong study design for key question.

Exclusion: wrong study design for key question.

Exclusion: wrong publication type.

Exclusion: wrong publication type.

Exclusion: wrong outcomes.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong comparison.

Exclusion: wrong study design for key question.

Exclusion: wrong population.

Exclusion: wrong study design for key question.

Exclusion: wrong study design for key question.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong population.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.
Appendix A4. List of Excluded Studies


Appendix A4. List of Excluded Studies

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong outcomes.

Exclusion: wrong publication type.

Exclusion: wrong comparison.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong intervention.

Exclusion: wrong population.

Exclusion: wrong population.

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: wrong study design for key question.

Exclusion: review not meeting inclusion criteria.

Exclusion: wrong intervention.

Exclusion: wrong intervention.

Exclusion: wrong population.
Appendix A4. List of Excluded Studies

Exclusion: wrong population.

Exclusion: wrong population.

Exclusion: wrong study design for key question.

Exclusion: wrong comparison.

Exclusion: wrong outcomes.

Exclusion: wrong comparison.

Exclusion: wrong study design for key question.

Exclusion: wrong study design for key question.

Exclusion: wrong study design for key question.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong outcomes.

Exclusion: wrong intervention.

Exclusion: wrong study design for key question.

Exclusion: studies outside of search dates.

Welch HG, Frankel BA. Likelihood that a woman with screen-detected breast cancer has had her "life saved" by that screening. Arch Intern Med. 2011;171(22):2043-6.
Exclusion: wrong study design for key question.
Appendix A4. List of Excluded Studies

**Exclusion:** wrong intervention.

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** works outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong intervention.

**Exclusion:** wrong publication type.

**Exclusion:** wrong population.

**Exclusion:** wrong comparison.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong intervention.

**Exclusion:** wrong intervention.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong intervention.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong comparison.

Appendix A4. List of Excluded Studies

**Exclusion:** wrong study design for key question.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong comparison.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong population.


**Exclusion:** wrong comparison.

Yaffe MJ, Barnes GT, Orton CG. Point/Counterpoint. Film mammography for breast cancer screening in younger women is no longer appropriate because of the demonstrated superiority of digital mammography for this age group. Med Phys. 2006;33(11):3979-82.

**Exclusion:** wrong publication type.


**Exclusion:** wrong publication type.


**Exclusion:** wrong intervention.


**Exclusion:** wrong study design for key question.


**Exclusion:** wrong intervention.


**Exclusion:** wrong intervention.


**Exclusion:** wrong outcomes.


**Exclusion:** wrong population.


**Exclusion:** no original data to include; publication or dataset with longer followup, more complete data, or same data already included in review.


**Exclusion:** wrong intervention.

Appendix A4. List of Excluded Studies

**Exclusion**: wrong outcomes.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong comparison.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong comparison.


**Exclusion**: wrong comparison.


**Exclusion**: studies outside of search dates.

Yee KM. Cancer: Screening mammo cuts incidence of late-stage cancer. 2014

**Exclusion**: wrong publication type.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong outcomes.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong comparison.


**Exclusion**: wrong intervention.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong publication type.


**Exclusion**: wrong study design for key question.


**Exclusion**: wrong population.


**Exclusion**: wrong population.


**Exclusion**: wrong population.

Screening for Breast Cancer 246 Pacific Northwest EPC
Appendix A4. List of Excluded Studies

**Exclusion:** wrong outcomes.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** wrong publication type.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong outcomes.

**Exclusion:** wrong comparison.

**Exclusion:** wrong intervention.

**Exclusion:** wrong comparison.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong study design for key question.

**Exclusion:** wrong publication type.

**Exclusion:** wrong outcomes.
Appendix A5. Quality Rating Criteria

Randomized, Controlled Trials (RCTs) and Cohort Studies\textsuperscript{1,2}

Criteria:

- Initial assembly of comparable groups:
  - for RCTs: adequate randomization, including first concealment and whether potential confounders were distributed equally among groups
  - for cohort studies: consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, crossovers, adherence, contamination)
- Important differential loss to followup or overall high loss to followup
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- Important outcomes considered
- Analysis: adjustment for potential confounders for cohort studies, or intention-to-treat analysis for RCTs.

Definition of ratings based on above criteria:

Good: Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (followup at least 80 percent); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; important outcomes are considered; and appropriate attention to confounders in analysis. In addition, for RCTs, intention to treat analysis is used.

Fair: Studies will be graded “fair” if any or all of the following problems occur, without the fatal flaws noted in the “poor” category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred in followup; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention to treat analysis is done for RCTs.

Poor: Studies will be graded “poor” if any of the following fatal flaws exists: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. For RCTs, intention to treat is lacking.

Case Control Studies\textsuperscript{1,2}

Criteria:

- Accurate ascertainment of cases
- Nonbiased selection of cases/controls with exclusion criteria applied equally to both
- Response rate
- Diagnostic testing procedures applied equally to each group
- Measurement of exposure accurate and applied equally to each group
- Appropriate attention to potential confounding variable
Appendix A5. Quality Rating Criteria

**Definition of ratings based on criteria above:**

**Good:** Appropriate ascertainment of cases and nonbiased selection of case and control participants; exclusion criteria applied equally to cases and controls; response rate equal to or greater than 80 percent; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to confounding variables.

**Fair:** Recent, relevant, without major apparent selection or diagnostic work-up bias but with response rate less than 80 percent or attention to some but not all important confounding variables.

**Poor:** Major selection or diagnostic work-up biases, response rates less than 50 percent, or inattention to confounding variables.

**Systematic Reviews**

**Criteria:**

- Search dates reported?
- Search methods reported?
- Comprehensive search?
- Inclusion criteria reported?
- Selection bias avoided?
- Validity criteria reported?
- Validity assessed appropriately?
- Methods used to combine studies reported?
- Findings combined appropriately?
- Conclusions supported by data?

**Definitions of ratings based on above criteria:**

**Good:** Meets all criteria: reports comprehensive and reproducible search methods and results; reports pre-defined criteria to select studies and reports reasons for excluding potentially relevant studies; adequately evaluates quality of included studies and incorporates assessments of quality when synthesizing data; reports methods for synthesizing data and uses appropriate methods to combine data qualitatively or quantitatively; conclusions supported by the evidence reviewed.

**Fair:** Studies will be graded fair if they fail to meet one or more of the above criteria, but the limitations are not judged as being major.

**Poor:** Studies will be graded poor if they have a major limitation in one or more of the above criteria.

**References**

Appendix A5. Quality Rating Criteria

Breast Cancer Surveillance Consortium

In 1994 the National Cancer Institute (NCI) established the Breast Cancer Surveillance Consortium (BCSC) to study breast cancer screening practices in the United States, with the recognition that results from controlled clinical trials of mammography may differ from the results of community screening practices. Each of the Consortium’s breast imaging registries collects population-based screening and diagnostic mammography data and links it to state or regional cancer registries. The following BCSC registries contributed data to this report: the Carolina Mammography Registry (North Carolina), Group Health Cooperative (Seattle Puget Sound region), New Hampshire Mammography Network, San Francisco Mammography Registry, and Vermont Breast Cancer Surveillance System. Mammography data are also linked to pathology databases, which include benign as well as malignant outcomes. A comparison of women represented in the BCSC against 2000 Census data shows that Consortium sites are located in counties that contain slightly more than 5 percent of the U.S. population, and represent the population in important sociodemographic respects.¹

Currently, the Consortium's database contains information on 10.7 million mammography examinations (including 2.6 million digital), 2.5 million women, and 130,000 breast cancer cases. Information on the distribution of key variables, mammographic data, characteristics of cases, and screening performance, among others, are detailed on the BCSC website: http://breastscreening.cancer.gov/data/. Data are pooled at a central Statistical Coordinating Center.

Registries and the Coordinating Center received institutional review board approval for active or passive consenting processes or a waiver of consent to enroll participants, link data, and perform analysis. All procedures were Health Insurance Portability and Accountability Act compliant, and registries and the coordinating center received a federal Certificate of Confidentiality and other protections for the identities of women, physicians, and facilities.

BCSC data include screening as well as diagnostic mammography. Screening mammography examinations are those designated as such by the ordering provider or radiologist performed more than 9 months after a previous imaging examination in women without a history of breast cancer or breast augmentation. Unilateral exams are excluded. Mammography information includes breast density, Breast Imaging Reporting And Data System (BI-RADS) assessment score, and recommendations for further work-up. In addition, prior to each mammography examination, a woman fills out a questionnaire that includes demographic and medical history information, including previous mammography information. Each screening mammography examination is given initial BI-RADS score based on the screening views only, which categorizes it as “positive” or “negative.” In our analysis, an initial score of 0, 4, 5, or 3 with a recommendation for immediate work-up is considered positive, whereas a score of 1, 2 or 3 without a recommendation for immediate work-up is negative.

In this report, BCSC data from 2003 to 2011 are included to examine the 1) frequency of recommendations for additional imaging and biopsy procedures resulting from positive screening mammography, 2) potential adverse effects of mammography screening, 3) incidence of ductal carcinoma in situ and invasive cancers detected by mammography screening; and 4) differences...
Appendix A6. Details of the Breast Cancer Surveillance Consortium

in outcomes between groups based on age, risk factors, and time since last mammography screening. Information for women under age 40 years or who have histories of breast augmentation or previous breast cancer diagnosis has been excluded.

References


Appendix A7. List of Peer Reviewers

**Expert Reviewers**

**Joann G. Elmore, MD, MPH** Professor of Medicine, Adjunct Professor of Epidemiology, University of Washington, Seattle, WA

**Laura Esserman, MD, MBA** Director, UCSF Carol Franc Buck Breast Care Center, Professor of Surgery and Radiology, University of California, San Francisco, CA

**W. Phil Evans, MD** Clinical Professor of Radiology; Director, UT Southwestern Center for Breast Care, University of Texas Southwestern Medical Center, Dallas, TX

**Suzanne W. Fletcher, MD, MSc** Professor emerita, Department of Population medicine, Harvard medical School and Harvard Pilgrim Health Care Institute, Cambridge, MA

**Karla Kerlikowske, MD** Professor of Medicine and Epidemiology and biostatistics, University of California, San Francisco, CA

**Anthony B. Miller, MD, FRCP, FRCP(C), FACE, FFPHM** Professor Emeritus, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

**Edward A. Sickles, MD** Professor Emeritus of Radiology, UCSF School of Medicine, San Francisco, CA

**Federal Reviewer**

**Barnett (Barry) Kramer, MD, MPH** Director, Division of Cancer prevention, National Cancer Institute, Rockville, MD
Appendix A8. List of Studies in the Included Systematic Reviews

Key Question 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?

**Broeders et al, 2012**


Appendix A8. List of Studies in the Included Systematic Reviews


Appendix A8. List of Studies in the Included Systematic Reviews


Harris et al, 2011


Appendix A8. List of Studies in the Included Systematic Reviews


Moss et al, 2012


Appendix A8. List of Studies in the Included Systematic Reviews


Njor et al, 2012


Appendix A8. List of Studies in the Included Systematic Reviews


Key Question 5. What are the harms of routine mammography screening, and how do they differ by age, risk factor, and screening interval?

Psychological Harms

Bond et al, 2013


Appendix A8. List of Studies in the Included Systematic Reviews

Hafslund and Nortvedt, 2009


Appendix A8. List of Studies in the Included Systematic Reviews

**Brewer et al, 2007**


Appendix A8. List of Studies in the Included Systematic Reviews


Appendix A8. List of Studies in the Included Systematic Reviews


Appendix A8. List of Studies in the Included Systematic Reviews


---

**Pain from Mammography**

**Whelehan et al, 2013**


Appendix A8. List of Studies in the Included Systematic Reviews


Armstrong et al, 2007

Appendix A8. List of Studies in the Included Systematic Reviews


---

Miller et al, 2008


Lambertz CK. Pre-medication with Acetaminophen for Screening Mammography: Gonzaga University; 1998.
Appendix A8. List of Studies in the Included Systematic Reviews


---

**Overdiagnosis**

**Marmot et al, 2013**


---

**Moss, 2005**


Appendix A8. List of Studies in the Included Systematic Reviews

Puliti et al., 2012


## Appendix B1. Quality Ratings of Systematic Reviews

<table>
<thead>
<tr>
<th>Author, year, title</th>
<th>Study design predetermined?</th>
<th>Dual review of studies and data abstraction?</th>
<th>Comprehensive literature search?</th>
<th>Publication status used as inclusion criteria?</th>
<th>List of included and excluded studies provided?</th>
<th>Characteristics of included studies provided?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armstrong et al., 2007</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Included studies: Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Bond et al, 2013</td>
<td>Yes</td>
<td>Dual review of abstracts</td>
<td>No</td>
<td>Included studies: Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Brett et al, 2005</td>
<td>No</td>
<td>Not reported</td>
<td>Yes</td>
<td>No</td>
<td>Included studies: No</td>
<td>Yes</td>
</tr>
<tr>
<td>Brewer et al, 2007</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Included studies: No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Broeders et al, 2012</td>
<td>Yes</td>
<td>Not reported</td>
<td>Yes</td>
<td>Included studies: No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Hafslund and Nortvedt, 2009</td>
<td>Unclear</td>
<td>Not reported</td>
<td>Yes</td>
<td>Included studies: No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Harris et al, 2011</td>
<td>Yes</td>
<td>Dual review of abstracts</td>
<td>No</td>
<td>Included studies: No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Marmot et al, 2013</td>
<td>Unclear</td>
<td>Not reported</td>
<td>Unclear</td>
<td>Included studies: Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Miller et al, 2008</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Included studies: Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Moss et al, 2012</td>
<td>Yes</td>
<td>Not reported</td>
<td>Yes</td>
<td>Included studies: No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Njor et al, 2012</td>
<td>Yes</td>
<td>Not reported</td>
<td>Yes</td>
<td>Included studies: No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Puliti et al, 2012</td>
<td>Yes</td>
<td>Not reported</td>
<td>Yes</td>
<td>Included studies: No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Whelehan et al, 2013</td>
<td>Yes</td>
<td>Dual review of abstracts</td>
<td>Unclear</td>
<td>Included studies: Yes</td>
<td>Not all studies</td>
<td></td>
</tr>
</tbody>
</table>

### Quality Ratings of Systematic Reviews

<table>
<thead>
<tr>
<th>Author, year, title</th>
<th>Included studies quality assessed?</th>
<th>Quality of included studies considered in formulating conclusions?</th>
<th>Appropriate methods used to combine studies?</th>
<th>Publication bias assessed?</th>
<th>Conflict of interest stated?</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armstrong et al, 2007</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>No</td>
<td>Yes</td>
<td>Good</td>
</tr>
<tr>
<td>Bond et al, 2013</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Good</td>
</tr>
<tr>
<td>Brett et al, 2005</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Fair</td>
</tr>
<tr>
<td>Brewer et al, 2007</td>
<td>Coded variables, but not formally assessed with criteria</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Broeders et al, 2012</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Good</td>
</tr>
<tr>
<td>Hafslund and Nortvedt, 2009</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Fair</td>
</tr>
</tbody>
</table>
## Appendix B1. Quality Ratings of Systematic Reviews

<table>
<thead>
<tr>
<th>Author, year, title</th>
<th>Included studies quality assessed?</th>
<th>Quality of included studies considered in formulating conclusions?</th>
<th>Appropriate methods used to combine studies?</th>
<th>Publication bias assessed?</th>
<th>Conflict of interest stated?</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harris et al, 2011</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Good</td>
</tr>
<tr>
<td>Marmot et al, 2013</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Fair</td>
</tr>
<tr>
<td>Miller et al, 2008</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Good</td>
</tr>
<tr>
<td>Moss et al, 2012</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Good</td>
</tr>
<tr>
<td>Njor et al, 2012</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Good</td>
</tr>
<tr>
<td>Puliti et al, 2012</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Whelehan et al, 2013</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Fair</td>
</tr>
</tbody>
</table>
## Appendix B2. Evidence Table of Results of New Studies of Psychological Harms

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Age, years (mean or %); setting; population</th>
<th>Comparisons</th>
<th>Measures</th>
<th>Outcomes</th>
<th>Quality rating; limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bredal et al, 2013</td>
<td>Before-after study</td>
<td>57.7; women recalled in a screening program in Oslo, Norway</td>
<td>FP (n=560) and TP (n=80) at recall vs. 4 weeks later</td>
<td>HADS (score ≥11)</td>
<td><strong>Recall vs. 4 weeks later:</strong> anxiety (% cases): FP 15% vs. 5.5% (NS), TP 19% vs. 16.7% (NS); depression (% cases): FP 1.4% vs. 1.3% (NS), TP 1.3% vs. 6.9% (NS). Factors predicting anxiety or depression in multivariate models: low general life expectations, previous history of anxiety and/or depression, anxiety at baseline, satisfaction with information (predicts depression only).</td>
<td>NA*</td>
</tr>
<tr>
<td>Brodersen and Siersma, 2013</td>
<td>Nested case-control</td>
<td>28% 50-54, 32% 55-59, 23% 60-64, 17% ≥65; women in screening programs in Copenhagen and Funen, Denmark; cases=recalled; controls=normal results in the same clinic and day as cases</td>
<td>FP (n=272) vs. Normal screen (n=864) vs. TP (n=174)</td>
<td>COS-BC</td>
<td><strong>After screening mammography:</strong> Normal screen vs. FP and TP had significantly better scores on subscales for sense of dejection, anxiety, negative impact on behavior, sleep, or sexuality, breast examination, and on single items of feeling less attractive and keeping mind off things (p&lt;0.001 for all outcomes); no differences between FP and TP on any subscales. <strong>3 year followup:</strong> TP vs. FP and Normal screen had significantly worse scores on subscales of sense of dejection, anxiety, negative impact on behavior, sleep, or sexuality, social network, and on single items of feeling less attractive and keeping mind off things (p&lt;0.001 for all outcomes) and additional differences vs. Normal screen on subscales of inner calm, social networking, and existential values (p&lt;0.001 for all outcomes); FP vs. Normal screen had significantly worse scores on subscales for sense of dejection, anxiety, negative impact on behavior, sleep, or sexuality, breast examination, inner calm, social network, existential values, and on single items of feeling less attractive and keeping mind off things (p&lt;0.05).</td>
<td>Good</td>
</tr>
<tr>
<td>Espasa et al, 2012</td>
<td>Case-control</td>
<td>55% 50-59, 45% 60-69; women in screening program in Spain; cases=FP; controls=TN matched on age, education, marital and working status, and previous mammograms</td>
<td>FP (n=100) vs. Normal screen (n=50)</td>
<td>HADS, structured interview</td>
<td><strong>After 22 days of followup:</strong> FP vs. Normal screen worried about having breast cancer (49% vs. 10%, p&lt;0.0001) and had worries that affected mood or daily activities (31% vs. 2%, p&lt;0.0001); but no differences in anxiety (11% vs. 14%, p=0.83) or depression (2% vs. 2%).</td>
<td>Fair; enrolled selected group of women; 2:1 ratio of cases to controls; did not control for confounders</td>
</tr>
</tbody>
</table>
# Appendix B2. Evidence Table of Results of New Studies of Psychological Harms

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Age, years (mean or %); setting; population</th>
<th>Comparisons</th>
<th>Measures</th>
<th>Outcomes</th>
<th>Quality rating; limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitzpatrick et al, 2011</td>
<td>Retrospective cohort</td>
<td>Mean age: NR, range: 50-62; women screened through the National Breast Screening Programme in Ireland</td>
<td>FP (n=9,746) vs. Normal screen (n=148,589)</td>
<td>Re-attendance</td>
<td>Rate of re-attendance: 90.7% vs. 89.0%, p&lt;0.001; age group 50-54 years: 91.0% vs. 89.6%, p&lt;0.001; age group 55-59 years: 90.4% vs. 88.7%, p&lt;0.001; age group 60-62 years: 90.4% vs. 87.4%, p&lt;0.001</td>
<td>Fair, unclear if random or consecutive sample; baseline data not provided; did not control for confounders</td>
</tr>
<tr>
<td>Gibson et al, 2009</td>
<td>Prospective cohort</td>
<td>6% &lt;50, 32% 50-59, 34% 60-69, 22% 70-79, 6% ≥80; women registered in the New Hampshire Mammography Network and the New Hampshire Women for Health study</td>
<td>FP (n=2,107) vs. Normal screen (n=11,384) reference group</td>
<td>WHQ</td>
<td>OR for depression (95% CI): overall FP 0.96 (0.72 to 1.28); white FP 0.84 (0.62 to 1.15); non-white FP 3.23 (1.32 to 7.91).</td>
<td>Fair; unclear how women were selected; baseline data not provided for groups of interest; outcomes self-reported</td>
</tr>
<tr>
<td>Hafslund et al, 2012</td>
<td>Nested case-control</td>
<td>57 (SD 5.8) for FP vs. 58 (SD 5.5) for TN; women from Hordaland, Sogn, and Fjordane Counties, Norway; cases=FP; controls=TN</td>
<td>FP (n=128) vs. Normal screen (n=195)</td>
<td>SF-36, HADS</td>
<td>6 months followup: FP vs. Normal screen clinical anxiety (mean HADS-A) 4.1 vs. 4.0, p=0.81; clinical depression (mean HADS-D) 3.2 vs. 2.4, p=0.045; mental function (mean SF-36) 80.6 vs. 85.0, p=0.03; vitality (mean SF-36) 70.3 vs. 77.0, p=0.02.</td>
<td>Fair; enrolled selected group of women; higher response rate in control group</td>
</tr>
<tr>
<td>Keyzer-Dekker, 2012</td>
<td>Prospective cohort</td>
<td>50 (SD 0.8) for 1st screen recalls vs. 61 (SD 5.9) for repeat screen recalls, p&lt;0.001; women with abnormal results referred to hospitals in The Netherlands</td>
<td>1st screen recalls (n=186) vs. repeat screen recalls (n=296)</td>
<td>STAI, NEO-FFI, CES-D, WHOQOL</td>
<td>After recall before diagnosis: anxiety (mean STAI) 13.3 vs. 12.8, p=0.209; depression (mean CES-D) 8.9 vs. 9.0, p=0.836.</td>
<td>Fair; older women in repeat screen group; outcomes were self-reported; did not report attrition</td>
</tr>
</tbody>
</table>
## Appendix B2. Evidence Table of Results of New Studies of Psychological Harms

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>Age, years (mean or %); setting; population</th>
<th>Comparisons</th>
<th>Measures</th>
<th>Outcomes</th>
<th>Quality rating; limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klompenhouwer et al, 2014&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Retrospective cohort</td>
<td>Mean age: NR, range: 50-75; women being screened in one of the specialized screening units in The Netherlands</td>
<td>Normal screen (n=373,474) vs. 1&lt;sup&gt;st&lt;/sup&gt; screen recalls (n=6,672) vs. repeat screen recalls for different lesion (n=161) vs. repeat screen recalls for same lesion (n=89)</td>
<td>Re-attendance rates</td>
<td><strong>Rate of re-attendance:</strong> 93.2% (95% CI, 93.1% to 93.3%) vs. 65.4% (95% CI, 64.0% to 66.8%) vs. 56.7% (95% CI, 47.1% to 66.4%) vs. 44.3% (95% CI, 31.4% to 57.1%); and 52.1% (95% CI, 44.4% to 59.8%) for all recalled groups combined</td>
<td>Fair, baseline data not provided for groups; did not control for confounders</td>
</tr>
<tr>
<td>Maxwell et al, 2013&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Retrospective cohort</td>
<td>Mean age: NR, range: 49-66; women screened at 1 of 5 breast screening programs in the United Kingdom</td>
<td>FP (n=9,367) vs. Normal screen (n=243,650) and Prevalent screen (n=54,716) vs. incident screen (n=198,301)</td>
<td>Re-attendance rates</td>
<td><strong>Rate of re-attendance:</strong> 87.7% of prevalent FP screen vs. 86.0% of prevalent normal screen, difference of 1.61% (95% CI, 0.54% to 2.62%); 92.0% of incident FP vs. 92.4% of incident normal screen, difference of -0.04% (95% CI, -1.18% to 0.31%); 86.2% of all prevalent screens vs. 92.4% of all incident screens</td>
<td>Fair, baseline data not provided for groups; did not control for confounders</td>
</tr>
<tr>
<td>Tosteson et al, 2014&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Nested case-control</td>
<td>41% &lt;50, 45% 50-64, 14% ≥65 years; women participating in DMIST in the United States; cases=FP, controls=TN matched by institution and age</td>
<td>FP (n=494) vs. Normal screen (n=534)</td>
<td>STAI, EuroQOL EQ-5D</td>
<td><strong>After mammography:</strong> FP vs. Normal screen anxiety (mean STAI) 35 vs. 33, p=NR; QOL (mean EQ-5D) 0.90 vs. 0.90, p=NR. <strong>1 year followup:</strong> FP anxiety (STAI mean difference) -1.53 (SD 13.14), p=0.01; QOL (EQ-5D mean difference) 0.001 (SD 0.13), p=0.13; Normal screen anxiety and QOL did not change over time.</td>
<td>Good</td>
</tr>
</tbody>
</table>

*Quality rating criteria not available for this study design.*
Appendix B2. Evidence Table of Results of New Studies of Psychological Harms

**Abbreviations:** CES-D=Center for Epidemiological Studies-Depression Scale; CI=confidence interval; COS-BC=Consequences of Screening in Breast Cancer; DMIST=Digital Mammographic Imaging Screening Trial; FP=false-positive; HADS=Hospital Depression and Anxiety Scale; HADS-A=HADS-Anxiety Subscale; HADS-D=HADS-Depression Subscale; n=number; NA=not available; NEO-FFI=Neuroticism-Extraversion-Openness-Five Factor Inventory; NR=not reported; NS=not significant; OR=odds ratio; QOL=quality of life; SD=standard deviation; SF-36=Short-form 36 Health Survey; STAI=Spielberger State-Trait Anxiety Inventory; TP=true positive; vs.=versus; WHOQOL=World Health Organization Quality of Life Assessment Instrument; WHQ=Women's Health Questionnaire.
### Appendix B3. Quality Ratings of Cohort Studies

| Author, Year              | Did the study attempt to enroll a random sample or consecutive patients meeting inclusion criteria (inception cohort)? | Were the groups comparable at baseline? | Did the study use accurate methods for ascertaining exposures, potential confounders, and outcomes? | Were outcome assessors and/or data analysts blinded to treatment? | Did the article report attrition? | Quality rating |
|---------------------------|-------------------------------------------------------------------------------------------------|----------------------------------------|-------------------------------------------------------------------------------------------------|----------------------------------------------------------------|
| Bredal et al, 2013107     | Yes                                                                                             | Unclear                                | Yes                                                                                             | No - self report                                                   | Yes                              | Fair           |
| Elmore et al, 1998149     | Yes                                                                                             | Yes                                    | Unclear                                                                                         | No                                                                | Not applicable                   | Fair           |
| Fitzpatrick et al, 2011200| Unclear                                                                                         | Unclear                                | Yes                                                                                             | No                                                                | Not applicable                   | Fair           |
| Gibson et al, 2009198     | Unclear                                                                                         | Unclear                                | Yes                                                                                             | No                                                                | Not applicable                   | Fair           |
| Hellquist et al, 2011123  | Yes                                                                                             | Mostly, ages women were included changed over time | Yes                                                                                             | No                                                                | Not applicable                   | Fair           |
| Hubbard et al, 2011143    | Yes                                                                                             | Yes                                    | Yes                                                                                             | No                                                                | Yes                              | Fair           |
| Kerlikowske et al, 2013144| Yes                                                                                             | Yes                                    | Yes                                                                                             | No                                                                | Yes                              | Fair           |
| Keyzer-Dekker et al, 2012159| Yes                                                                                         | Mostly, older women in repeat screen group (p<0.001) | Yes                                                                                             | No - self report                                                   | No                               | Fair           |
| Klompenhouwer et al, 2014202| Yes                                                                                         | Unclear                                | Yes                                                                                             | No                                                                | Not applicable                   | Fair           |
| Maxwell et al, 2013201    | Yes                                                                                             | Unclear                                | Yes                                                                                             | No                                                                | Not applicable                   | Fair           |

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Did the study perform appropriate statistical analyses on potential confounders?</th>
<th>Is there important differential loss to followup or overall high loss to followup?</th>
<th>Were outcomes prespecified and defined, and ascertained using accurate methods?</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bredal et al, 2013107</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Elmore et al, 1998149</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Fitzpatrick et al, 2011200</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Gibson et al, 2009198</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Hellquist et al, 2011123</td>
<td>Yes</td>
<td>Yes (5.5% of women had 1 year of followup and 2.9% observed for 10 or more years)</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Hubbard et al, 2011143</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Fair</td>
</tr>
</tbody>
</table>
## Appendix B3. Quality Ratings of Cohort Studies

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Did the study perform appropriate statistical analyses on potential confounders?</th>
<th>Is there important differential loss to followup or overall high loss to followup?</th>
<th>Were outcomes prespecified and defined, and ascertained using accurate methods?</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kerlikowske et al, 2013¹⁴⁴</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Keyzer-Dekker et al, 2012¹⁹⁹</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Klompenhouwer et al, 2014²⁰²</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Maxwell et al, 2013²⁰¹</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Yes</td>
<td>Fair</td>
</tr>
</tbody>
</table>
## Appendix B4. Quality Ratings for Case-Control Studies

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Did the study attempt to enroll all or random sample of cases using pre-defined criteria?</th>
<th>Were the controls derived from the same population as the cases?</th>
<th>Were the groups comparable at baseline on key prognostic factors?</th>
<th>Were enrollment rates similar in cases and controls invited to participate?</th>
<th>Did the study use accurate methods for identifying outcomes?</th>
<th>Did the study use accurate methods for ascertaining exposures and potential confounders?</th>
<th>Did the study perform appropriate statistical analyses on potential confounders?</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brodersen and Siersma, 2013 \textsuperscript{203}</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes, age is different, but would be expected</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Good</td>
</tr>
<tr>
<td>Espasa et al, 2012 \textsuperscript{206}</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes, matched</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Fair</td>
</tr>
<tr>
<td>Hafslund et al, 2012 \textsuperscript{205}</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>No, more controls responded (85% vs. 52%)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Fair</td>
</tr>
<tr>
<td>Tosteson et al, 2014 \textsuperscript{204}</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes, age is different, but would be expected</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Good</td>
</tr>
</tbody>
</table>