Breast Cancer Screening

October 7, 2011

Donna Fitzpatrick-Lewis, Nicole Hodgson, Donna Ciliska, Leslea Peirson, Mary Gauld, Yan Yun Liu
McMaster University
Hamilton, Ontario, Canada

CTFPHC Leads:
Marcello Tonelli and Richard Birtwhistle

PHAC Scientific Officer:
Sarah Connor Gorber

Working Task Force Members:
James Dickinson, Michel Joffres, Gabriela Lewin, Harminder Singh

Task Force Members:
Neil Bell, Paula Brauer, Martin Fortin, Patrice Lindsay,
Patricia Parkin, Kevin Pottie, Elizabeth Shaw
Abstract

Background: This systematic review is an update of the evidence since the 2001 Canadian Task Force recommendations on breast cancer screening.

Purpose: A decision was made to update the United States Preventive Services Task Force (USPSTF) 2009 review; therefore the purpose was the same as that review: to determine the effectiveness of mammography screening on decreasing breast cancer and all cause mortality among average-risk women between the ages of 40 and 49 and those 70 years and older; to determine the effectiveness of clinical breast examination (CBE) or breast self-examination (BSE) in decreasing mortality in average-risk women of the same age groups; and to determine the harms associated with mammography, CBE, and BSE. Additional contextual questions considered the costs associated with screening; patient preferences and values regarding breast screening; and particular subgroup information regarding the burden of breast cancer or rates of screening among Aboriginal women, rural or remote-residing women, and women of various ethnic backgrounds; and the optimal frequency of screening.

Data Sources: The search strategy from the USPSTF’s 2009 review of breast cancer screening was updated. Medline® and the Cochrane Database of Systematic Reviews were searched from December 2008 to October 2010 for studies in English and French. For patient preferences and values CINAHL and Medline were searched from 2000 to October 2010. Medline was searched back to 1950 for systematic reviews for subgroups. References of retrieved articles were checked, selected grey literature was searched for Canadian statistics, and some authors were contacted.

Study Selection: Randomized controlled trials and systematic reviews with breast cancer mortality or all cause mortality as outcomes for effectiveness of screening (mammography, CBE, or BSE) were included. For the literature on harms and on patient preferences and values, all study designs were included; for subgroups of interest, systematic reviews were included.

Data Abstraction: Relevant articles were abstracted. Study quality was assessed using GRADE.

Results: No new trials were found regarding mammography, CBE, or BSE on breast cancer mortality or all cause mortality. Seventeen new publications were identified and included: one systematic review of the effect of mammography on mortality; two systematic reviews and nine primary studies of harms; and five papers on costs. The search for information on patient preferences and values found three systematic reviews and 23 primary studies.

Data Synthesis: There were no new trials of mammography on breast cancer mortality; trials identified during the USPSTF search were summarized using the GRADE process. Of nine available trials, four were adequately randomized and five had methodological or reporting deficits related to randomization. In a meta-analysis of the eight studies (348,219 participants) of screening mammography in women aged 39–49 the pooled effect of screening versus no screening on breast cancer mortality was a relative risk (RR) of 0.85 (95% CI 0.75–0.96; $I^2=0\%$) after an overall median follow-up of 11.4 years. Pooled results from two trials showed that screening did not significantly reduce all cause mortality among 211,270 women aged 39–49 (RR 0.97, 95% CI 0.91–1.04; $I^2=0\%$). Meta-analysis of the two trials (17,646 participants) that reported results for women ≥70 years found that screening led to a nonsignificant reduction in breast cancer mortality (RR 0.68, 95% CI 0.45–1.01; $I^2=0\%$). The meta-analysis of seven studies
of mammography screening for the 250,274 women aged 50–69 years found a reduction in breast cancer mortality (RR 0.79, 95% CI 0.68–0.90; \(I^2=41\%\)). Although the relative risk reductions were statistically significant for most age groups, the absolute magnitude of the reductions was small across all age groups.

The effectiveness of BSE and CBE has not been established. Two studies of BSE from the USPSTF review showed no difference in breast cancer mortality.

Harms include false-positive rates of 6.5% for mammography and 8.7% for CBE and/or mammography. Approximately 28% of women aged 50-69 and 33% of women aged 40-49 screened with mammography will receive at least one false-positive result.

The studies of patient preferences and values found that women value mammography for the perceived reduction in mortality; few women consider issues of harm arising from false-positives in making decisions about breast cancer screening. Aboriginal women and women who live in rural and remote geographies have less access to mammography services than do women in the general population.

**Limitations:** The search was updated based on the USPSTF review; therefore, EMBASE was not searched, and only articles in English and French were included. The searches for cost, patient values and preferences, and special populations were focused and not based on a full systematic review.

**Conclusions:** This updated meta-analysis of mammography screening trials indicates a reduction in breast cancer mortality for women aged 40–49 and a nonsignificant effect on breast cancer mortality for women \(\geq 70\) years. Pooled analyses confirm the previously reported reduction in breast cancer mortality for women aged 50–69 years. Future trials will be essential in assessing risk and benefit in screening the Canadian population and in determining the effect of newer technologies for breast imaging.
Addendum


The literature search for systematic reviews on the effectiveness of screening for breast cancer was completed to October 2010. Prior to the review being posted, we became aware of new follow-up data published on the Swedish Two-County Trial (East County: Östergötland; West County: Kopparberg/Dalarna).\(^1\) This paper provides the longest follow-up period of all the breast cancer screening trials with 29-year follow-up data and gives estimates of the impact on mortality in both relative and absolute effects. The Swedish Two-County Trial was an RCT in which 133,065 women aged 40-74 were randomized to an invited to mammography screening group or to a control group that received usual care. The study authors provided and analyzed data on mortality based on two different committees that used slightly different inclusion criteria. In two separate analyses of breast cancer mortality for all ages, we pooled the most recent data from the Swedish Two-County Trial with data from the seven other screening trials included in our review. Our findings indicate that when the 29-year follow-up data are used there is a slight overall effect, moving the relative risk (RR) from 0.82 (95% CI 0.74-0.91) (see Forest Plot 1) to RR 0.81 (95% CI 0.74-0.88) with local trial end point committee data (see Forest Plot 2) or RR 0.81 (95% CI 0.75-0.88) with Swedish overview committee consensus data (see Forest Plot 3).
Table of Contents

Abstract...........................................................................................................................................i

Table of Contents..............................................................................................................................iii
  List of Tables.................................................................................................................................iv
  List of Figures.................................................................................................................................iv

Abbreviations.....................................................................................................................................v

Chapter Introduction.......................................................................................................................1
  Purpose...........................................................................................................................................1
  Condition Background....................................................................................................................1
  Risk Factors....................................................................................................................................2
  Rationale for Screening....................................................................................................................2
  Screening Strategies.......................................................................................................................3
  Interventions and Treatments..........................................................................................................3
  Current Clinical Practice..................................................................................................................4
  Previous Review and CTFPHC Recommendations.........................................................................4

Chapter 2: Methods.........................................................................................................................6
  Analytic Framework and Key Questions..........................................................................................6
  Search Strategies.............................................................................................................................6
  Study Selection.................................................................................................................................7
  External Review...............................................................................................................................7
  Quality Assessment, Data Abstraction, and Analysis.....................................................................7

Chapter 3: Results.............................................................................................................................10
  Summary of the Literature Search....................................................................................................10
  Results for Key Questions...............................................................................................................10
  Results for Contextual Questions....................................................................................................20

Discussion.........................................................................................................................................28
  Limitations.......................................................................................................................................29
  Future Research...............................................................................................................................30

Conclusion.........................................................................................................................................31
Reference List

List of Tables
Table 1: Harms from Screening – Ranking of Importance to Decision Making
Table 2: Relative Risks for Breast Cancer Mortality from Mammography Trials for All Ages
Table 3: Summary of Evidence – Mammography Trials for Women Younger than 50 Years – Breast Cancer Mortality
Table 4: Summary of Evidence – Mammography Trials for Women Younger than 50 Years – All Cause Mortality
Table 5: Estimated number of women with adverse outcomes following screening mammography
Table 6: Psychological Effects of False-Positive Mammograms
Table 7: Relative Risk of Breast Cancer Mortality for Mammography Screening Intervals

List of Figures
Figure 1: Analytic Framework and Key Questions
Figure 2: Search Results
Figure 3: Preference and Values Search Results

Appendices
Appendix 1: Search Terms for Mammography, Harms, and Costs
Appendix 2: Search Terms for Patient Preferences and Values
Appendix 3: Search Terms for Subpopulations
Appendix 4: Search Terms for Breast Cancer Frequency
Appendix 5: Grey Literature Search
Appendix 6: Characteristics of Included Studies
Appendix 7: Evidence Sets
Appendix 8: List of Studies Excluded at Full Text Screening
Appendix 9: List of External Reviewers – Protocol
Appendix 10: List of External Reviewers – Evidence Synthesis

Acknowledgements
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADH</td>
<td>atypical ductal hyperplasia</td>
</tr>
<tr>
<td>BSE</td>
<td>breast self-examination</td>
</tr>
<tr>
<td>CBE</td>
<td>clinical breast examination</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CINAHL</td>
<td>Cumulative Index to Nursing and Allied Health Literature</td>
</tr>
<tr>
<td>CNBSS</td>
<td>Canadian National Breast Screening Study</td>
</tr>
<tr>
<td>CTFPHC</td>
<td>Canadian Task Force on Preventive Health Care</td>
</tr>
<tr>
<td>DCIS</td>
<td>ductal carcinoma in situ</td>
</tr>
<tr>
<td>GRADE</td>
<td>Grading of Recommendations Assessment, Development and Evaluation</td>
</tr>
<tr>
<td>LCIS</td>
<td>lobular carcinoma in situ</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
<tr>
<td>OR</td>
<td>odds ratio</td>
</tr>
<tr>
<td>RCT</td>
<td>randomized controlled trial</td>
</tr>
<tr>
<td>RR</td>
<td>relative risk</td>
</tr>
<tr>
<td>UKBSP</td>
<td>United Kingdom Breast Screening Programme</td>
</tr>
<tr>
<td>USPSTF</td>
<td>United States Preventive Services Task Force</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

Purpose

The purpose of this review is to provide an update of the 2001 Canadian Task Force recommendations on breast cancer screening.\textsuperscript{1,2} A decision was made to update the United States Preventive Services Task Force (USPSTF) review of 2009.\textsuperscript{3} Previous reviews found fair evidence for mammography every one to two years for women 40 years and older; with benefits increasing and harms decreasing with age.\textsuperscript{3} Identified gaps included estimates of the proportion of benefits due to screening and cost-effectiveness of screening before age 50 and after age 69.\textsuperscript{2,3}

The goal of this review is to determine the effectiveness and harms of mammography screening, clinical breast examination (CBE) and breast self-examination (BSE) among average-risk women aged 40 to 49 years and 70 years and older. Comparison data for women aged 50 to 69 are included.

The USPSTF updated its 2002 guidelines in 2009.\textsuperscript{3,4} The 2009 update had key differences compared to the 2002 guidelines in terms of the recommendations for different age groups. The absence of current Canadian recommendations and the differences between the 2002 and 2009 USPSTF recommendations were the basis for selecting this topic for an update by the revitalized Canadian Task Force on Preventive Health Care (CTFPHC) in 2010. Due to the update nature of this review, some of both the background and methods sections rely on the USPSTF report.\textsuperscript{3}

Condition Background

Definition

Breast cancer consists of malignant cells resulting in a continuum from noninvasive to invasive carcinomas.\textsuperscript{3,5} The most common form is ductal carcinoma; there are a number of other subtypes of noninvasive and invasive lesions.

Noninvasive carcinomas are a proliferation of the mammary duct epithelial cells [ductal carcinoma in situ (DCIS)], or of the lobule [lobular carcinoma in situ (LCIS)]. Noninvasive lesions do not metastasize; however, DCIS, with its several subtypes,\textsuperscript{3,6} is considered to be a precursor for invasive ductal carcinoma, while LCIS (a bystander lesion found incidentally on biopsies) is considered to be a marker for increased risk of ductal or lobular cancer.\textsuperscript{6}

Invasive lesions have metastatic potential as they invade the basement membrane into the stroma. Metastatic breast cancer or secondary cancer means that the cancer has spread to other sites in the body, most commonly adjacent lymph nodes, bone, liver, lung, and brain.\textsuperscript{3,5} Breast cancer that has spread to other parts of the body is not curable; however, women with metastatic breast cancer can be treated.\textsuperscript{7} Approximately 70\% to 80\% of invasive breast cancers are invasive ductal carcinoma and 10\% are invasive lobular carcinoma; the remainder are special types (e.g., mucinous, tubular, adenoid cystic, etc.).\textsuperscript{3,5}

Prevalence and burden of disease

While breast cancer can occur in men, incidence is higher in women. For women aged 20 to 59 years, breast cancer is the second most common form of cancer and most common cause of cancer death.\textsuperscript{8} In 2010, approximately 28\% of new breast cancer cases diagnosed in Canada were in
women over age 69, and approximately 20% were in women under 50 years. There is little variation by province. 

**Etiology and natural history**

Breast cancer development is attributed to dysfunction in cell cycle regulation. Inherited and acquired mutations may influence the cycle. The majority of breast cancers are sporadic (over 90%) and unrelated to family history. Approximately 5% to 10% of breast cancers can be attributed to mutations in the genes BRCA1 and BRCA2, but other genes have also been studied.

Environmental exposures to hormones, diet, and viruses may play a role, but no single factor has been isolated. The precise role these factors play in tumor development is not clear, but each factor may be responsible for different steps of a series required to create malignant cells.

Information about the natural history of DCIS is lacking because it historically was treated by mastectomy. Ductal carcinoma in situ can recur or progress to invasive breast cancer, which has led to two conflicting models to explain the relationship between DCIS and invasive cancer: parallel disease and linear progression. The relationship is probably more complex, and both models may occur simultaneously. This co-occurrence is supported by different studies, including immunohistochemical analysis and gene expression profiling.

**Consequences if left untreated**

Different types of breast cancer have different growth rates, dependent on tumor biology. There are few reports of untreated patients; however, poor survival is characteristic of locally advanced breast cancer. Erbas and colleagues reviewed studies where DCIS was initially misdiagnosed as benign and treated by biopsy alone; 14% to 53% of patients with DCIS progressed to a diagnosis of invasive cancer over a period of 10 or more years.

**Risk Factors**

The most important risk factors for breast cancer are gender and age: 80% of all new breast cancer diagnoses are in women over the age of 50. Risk factors for invasive cancer include a history of noninvasive breast cancer or previous abnormal biopsy containing LCIS or atypical ductal hyperplasia (ADH). Strength of family history as a risk factor for breast cancer is related to the number of relatives affected, the degree of the relationships, and age at diagnosis of the family members.

Early age at menarche, older age at menopause, postmenopausal hormone replacement therapy, and postmenopausal obesity are all associated with increased risk for breast cancer. Other risk factors such as environmental exposures to radiation, therapeutic radiation (commonly given for lymphoma), and excess alcohol intake have been documented.

**Rationale for Screening**

There is widespread acceptance of the value of regular breast cancer screening as the single most important public health strategy to reduce breast cancer mortality.

Mammography, CBE, and BSE can all identify tumors. Mammography can identify asymptomatic breast cancer. Breast cancer can be more effectively treated at the asymptomatic stage. A recent
systematic review concluded that mammography screening is likely to reduce breast cancer mortality by an estimated 15%, corresponding with an absolute risk reduction of 0.05%.\textsuperscript{18}

**Screening Strategies**

The screening strategies considered in this review are mammography, BSE, and CBE. The USPSTF 2002 review found mammography screening is sensitive (77\%–95\% for all ages), but with lower sensitivity for women under age 50 (58\%–85\%); it is specific (94\%–97\%) and acceptable to most women.\textsuperscript{4} Mammography is in the process of shifting from film to digital technologies.

Both BSE and CBE have been promoted as inexpensive screening strategies. Breast self-exam has been suggested as a monthly examination of the woman’s breasts. There are varying estimates of the sensitivity (12\%–41\%),\textsuperscript{4} specificity has been estimated between 66\% and 81\%.\textsuperscript{19} In a review of 20 observational studies and three trials, Hackshaw and Paul\textsuperscript{20} concluded that regular BSE was not an effective method of reducing breast cancer mortality as there was no difference in mortality for those who had detected their cancer during a self-examination or for those who reported practicing BSE regularly. However BSE was associated with more women seeking medical care and having biopsies.\textsuperscript{20}

Clinical breast exam is the examination of the breasts by a health professional. Effectiveness of examination of the breasts by clinicians is highly influenced by the training and skills of the practitioner, age of the woman, and tumor size. CBE “sensitivity ranges from 40\% to 69\%, specificity from 88\% to 99\%, and positive predictive value from 4\% to 50\%.”\textsuperscript{21} (p. E354)

Positive outcomes of breast cancer screening must be put into the context of costs to the individual and to the healthcare system, considering benefits of tumour detection and earlier treatment, with emotional costs to patients and families due to false positive results and additional diagnostic tests and surgeries. One review found that screening led to up to 30\% overdiagnosis and overtreatment.\textsuperscript{18} In this context, overdiagnosis is defined as detection of invasive or noninvasive breast cancer that would not have been identified clinically or resulted in symptoms or death in a person’s lifetime. One of two large trials comparing BSE with no intervention found increased detection of tumors, but neither study found differences in breast cancer mortality.\textsuperscript{22} Few studies have assessed CBE.\textsuperscript{3}

Mammography screening is widely available in urban areas in Canada, with some mobile clinics for more rural areas. Cost calculations must consider the overall program cost, cost per screening exam, cost per cancer detected, and ultimately, overall cost-effectiveness, as measured by the cost per year of life gained. In 1996/97 in British Columbia, the total provincial costs for mammography screening were approximately $14 million, the cost per screening exam was $45.94, and the cost per cancer detected was $15,211.\textsuperscript{23}

**Interventions and Treatments**

Women with positive findings on BSE, CBE, or mammography are advised to undergo additional diagnostic tests, which may include further mammography, ultrasound, magnetic resonance imaging (MRI), and/or tissue sampling via needle core biopsy. Tissue testing includes identification of tumor type and preliminary grade, as well as examination of cellular receptors.\textsuperscript{3}
The goal of therapy is to improve survival, reduce recurrence, delay disease progression, maximize the patient’s quality of life, and support the patient and family. Treatment usually requires combinations of therapies, including surgery, chemotherapy, hormonal therapy, and radiation, depending on type and stage of cancer.3

**Current Clinical Practice**

In Canada, several guidelines recommend that women aged 50 years and older have a screening mammogram every two years, and that women aged 40 to 49 years talk to their healthcare providers to make personal decisions about mammography.24,25 In 2008, 72% of women aged 50 to 69 self-reported having had a mammogram in the past two years.26 However, epidemiological evidence indicates that in Canada the target participation rate of 70% in organized screening programs has not been reached.24,27

Abdel-Malek and colleagues conducted a cross-sectional study of general and family physicians active in Ontario.28 Adherence to screening was defined as recommending screening every two years to women aged 50 to 69 years. Only 38.9% of physicians followed recommended breast screening guidelines. After adjusting for physician gender and age, predictors of screening adherence included physicians working in academic or research centres (odds ratio [OR] 8.3, 95% CI 1.7–39.7), and those reporting that over 31% of their patients were of low income (OR 1.6, 95% CI 1.1–2.4). Those physicians located in a large city (>100,000 people), versus a rural area or town (<10,000 people), were less likely to adhere to screening guidelines (OR 0.5, 95% CI 0.3–0.7).

**Previous Review and CTFPHC Recommendations**

In 1994, the Canadian Task Force on the Periodic Health Exam published a guideline on breast cancer screening.29 In 2001, it was updated in two separate publications: recommendations for screening mammography among women aged 40–49 years at average-risk of breast cancer,2 and routine teaching of BSE for breast cancer.1 The first concluded that the evidence did not support inclusion or exclusion of screening mammography for women aged 40–49 years at average-risk of breast cancer (Grade C recommendation).2 With regard to teaching women BSE to screen for breast cancer1:

- women aged 40–49 and 50–69 years – it was recommended that routine teaching of BSE be excluded from the periodic health exam (Grade D recommendation)
- women aged <40 years and ≥70 years – there was insufficient evidence to make a recommendation.

In 2002, the USPSTF recommended mammography screening, with or without CBE, every one to two years for women aged 40 years and older.30 It concluded that evidence was insufficient to recommend for or against routine CBE alone and for or against teaching or performing routine BSE.30 The 2009 update found3,31:

- “Mammography screening reduces breast cancer mortality by 15% for women aged 39–49 (RR 0.85, 95% CI 0.75–0.96); data are lacking for women 70 years and older.
- Radiation exposure from mammography is low.
- Adverse experiences are common and transient.
- Estimates of overdiagnosis vary from 1% to 10%.
• Younger women have more false-positive results and additional imaging but fewer biopsies than older women.”3 (p. iii)

The absence of current Canadian recommendations and the differences between the 2002 and 2009 USPSTF recommendations were the basis for selecting this topic for an update by the revitalized Canadian Task Force on Preventive Health Care (CTFPHC) in 2010.3,30
Chapter 2: Methods

Analytic Framework and Key Questions

The analytic framework and key questions for this review follow the USPSTF questions for the 2009 update (Figure 1).³ The population of interest includes average-risk women; that is, women without pre-existing disease and those not considered to be at risk based on family history of breast or ovarian cancer or genetic mutations or abnormal breast pathology. As in the USPSTF report, the key questions focus on ages 40 to 49 years and over 69 years. However, data were extracted for the 50 to 69 year group as well. Key questions include:

1a. Does screening with mammography (film and digital) or MRI decrease breast cancer mortality and all cause mortality for women aged 40–49 and ≥70?

1b. Does CBE screening decrease breast cancer mortality for women aged 40–49 and ≥70? Alone or with mammography?

1c. Does BSE practice decrease breast cancer mortality for women aged ≥40?

2a. What are the harms associated with screening with mammography (film and digital) and MRI?

2b. What are the harms associated with CBE?

2c. What are the harms associated with BSE?

Additional contextual questions include:

1. What is the cost-effectiveness of screening?

2. What are patient preferences and values related to screening for breast cancer?

3. Are there subgroups of the Canadian population who have a higher prevalence of breast cancer or for whom it would be difficult to implement screening programs? Subgroup analyses that explore issues of burden of disease, screening rates, and special implementation issues include:
   i. Aboriginal
   ii. rural or remote-dwelling populations
   iii. ethnicity

4. What is the evidence of optimal frequency of screening with mammography?

Search Strategies

The USPSTF searched The Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews (through the fourth quarter of 2008), Medline® (January 2001 to December 1, 2008), reference lists and Web of Science for published studies, and the Breast Cancer Surveillance Consortium for screening mammography data.³ There were separate searches for screening, digital mammography, MRI, DCIS, adverse effects, and costs. For this update, the same search terms and databases were used, and all searches were updated to October 15, 2010. One search strategy was altered: the limits on study methods were removed in Medline, allowing
randomized controlled trials, meta-analyses, and systematic reviews to be left in the search. Reference lists of key articles were reviewed.

The EMBASE database was not searched, as it was not searched in the original review. An additional search was conducted to discover patient preferences and values; the Cumulative Index to Nursing and Allied Health Literature (CINAHL) and Medline were searched from 2000 to October 2010. Also, a search was done for particular subgroups including rural and remote populations, Aboriginal populations, and ethnic subgroups. Medline was searched for reviews (in English) back to 1950. Medline was searched for screening frequency. A specific search of the grey literature was also done in order to find relevant Canadian statistics, using the search terms “breast cancer screening AND harms”; “mammography AND harms”; “mammography AND costs”; and “breast cancer screening AND costs”. The detailed strategies for all searches are found in Appendices 1 through 5.

Study Selection

Eligible studies included women aged 40 years and older, without pre-existing breast cancer and not considered to be at high-risk for breast cancer on the basis of family history of breast or ovarian cancer or other personal risk factors, such as abnormal breast pathology or deleterious genetic mutations.

Study designs for effectiveness of screening (mammography, CBE, or BSE) included randomized controlled trials or meta-analyses with breast cancer mortality or all cause mortality as outcomes. For harms, studies of various designs and multiple data sources were included. Harms included radiation exposure, pain during procedures, patient anxiety and other psychological responses, consequences of false-positive and false-negative test results, and overdiagnosis.

Studies of cost-effectiveness of screening were included if they were relevant to the key questions. As was done for the USPSTF report, we excluded studies of costs of improving screening rates, dual review of screening mammography, or studies in populations at high-risk for breast cancer. Studies of patient preferences and values could be any study design, including qualitative studies. Studies of particular subgroups were systematic reviews. All included studies were in either English or French. Grey literature was included if it included recent relevant national Canadian data.

External Review

Before the review began, the protocol was internally reviewed by the Breast Cancer Working Group, which includes members of the CTFPHC and Public Health Agency of Canada staff. The revised protocol was sent to five external reviewers with expertise in review methodology and/or cancer; feedback was received from four reviewers of the protocol (Appendix 9), and revisions were made. A draft of the evidence review went to the Breast Cancer Working Group, and then the revised review went to external experts (Appendix 10) not affiliated with the CTFPHC.

Quality Assessment, Data Abstraction, and Analysis

The titles and abstracts were reviewed in duplicate by members of the synthesis team; any article marked for inclusion by either team member went on to full text rating. Full text inclusion, quality assessment, and data abstraction were done by two people. All disagreements were resolved through discussions rather than relying on a particular level of kappa score to indicate when discussions were no longer necessary. The inclusion results were reviewed by a third person. Data were
abstracted by two people using a standard format. The exception to this process was studies related to the contextual questions of costs, patient preferences, subpopulations, and grey literature, for which abstraction was done by one person.

The strength of the evidence was determined based on the GRADE system of rating quality of evidence using GRADEPro software. This system of grading evidence has been widely used and has been endorsed by more than 40 major organizations including the World Health Organization, Centers for Disease Control and Prevention, and the Agency for Healthcare Research and Quality. The GRADE system classifies quality of evidence in one of four levels: high, moderate, low, and very low. The final grade is based on: risk of bias due to limitations in design, inconsistency of findings, indirectness, imprecision, and publication bias.

The Breast Cancer Screening Working Group rated each of the outcomes and potential harms of screening using the GRADE process, which suggests a 9-point scale (1 through 9) to judge their importance. The upper end of the scale, rankings 7 through 9, identifies outcomes of critical importance for decision making. Rankings 4 through 6 represent outcomes that are important but not critical, while rankings 1 through 3 are items that are deemed to be of limited importance to decision making or to patients. This process identified breast cancer mortality and all cause mortality as the most important primary outcomes. The secondary outcomes of harms associated with screening were ranked as follows (Table 1).

**Table 1: Harms from Screening – Ranking of Importance to Decision Making**

<table>
<thead>
<tr>
<th>Harm</th>
<th>Importance</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>False-positive and false-negative mammography results, additional imaging and biopsies</td>
<td>Critical</td>
<td>7</td>
</tr>
<tr>
<td>Anxiety, distress, and other psychological responses</td>
<td>Important</td>
<td>6</td>
</tr>
<tr>
<td>Radiation exposure</td>
<td>Important</td>
<td>5</td>
</tr>
<tr>
<td>Overdiagnosis</td>
<td>Important</td>
<td>4.75</td>
</tr>
<tr>
<td>Pain during procedure</td>
<td>not important</td>
<td>3</td>
</tr>
</tbody>
</table>

The GRADE process was also used to assess risk of bias for individual studies, which was then used with the summary of findings to assess the overall quality of the evidence. In addition to those required data, for each study we abstracted data about the patient population, the study design, analysis, and results. Reviews were quality assessed using the assessment of multiple systematic reviews (AMSTAR) tool.

*GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.
Information to determine the quality of evidence was abstracted in duplicate from the primary methodology paper from each study. Those abstracting the data were blind to each other’s ratings. In cases of disagreement, final decisions were determined by consensus after consultation with a third reviewer. Separate tables were constructed and GRADE assessments were made by study design. When the method of randomization either was deemed inadequate (e.g., randomization by date of birth) or was not clear from the primary methodology publication, a separate table was constructed for RCTs and quasi-randomized trials. If the summary effect size from these subgroups of trials was similar and heterogeneity did not exist, the recommendations were based on all trials (i.e., randomized and quasi-randomized); otherwise, recommendations were based on the RCT results alone. In the first circumstance, it was reasoned that, although there was potential for bias due to inadequate randomization, evidence of this bias did not exist and therefore the overall estimate was the best estimate on which to base recommendations.

Traditional meta-analyses were undertaken using a random effects model proposed by DerSimonian and Laird.36 The random effects model assumes that the studies are a sample of all potential studies and incorporates an additional between-study component into the estimate of variability.

We used a test based on the deviations of the individual study estimates from the summary estimate of effect (the heterogeneity Chi²) as our primary method to test for heterogeneity.37 To supplement this test we calculated a statistic to quantify heterogeneity, the I², which describes the proportion of the variance in the point estimate due to heterogeneity rather than sampling error.38 Although there are no strict rules for interpreting I², a rough guide is that an I² greater than 50% may represent substantial heterogeneity.39

Publication bias was assessed using funnel plots, which graph the estimated effects against sample size. Funnel plot asymmetry indicates the likely presence of publication bias. However, there were at most nine trials included in any funnel plot, and the Cochrane Handbook suggests no fewer than ten trials,40 so these funnel plots are not included in this report, and we cannot be certain that publication bias is absent.
Chapter 3: Results

Summary of the Literature Search

The literature search for this review replicated and updated the search conducted by the USPSTF review in 2009. Our search located 920 potentially relevant citations (Figure 2). At title and abstract screening 700 were excluded. A total of 220 papers were retrieved and were assessed on inclusion criteria. Reasons for excluding at this level were: 105 because the population was high-risk; 6 because the intervention was not mammography, CBE, or BSE; 53 because the outcome studied was not mortality, costs, or harms; and 39 because the study design was not RCT or systematic review (with or without meta-analysis) for mortality. Seventeen papers met all the criteria. The papers, consisting of one systematic review of mortality, two systematic reviews of harms, and nine primary studies on harms have been included in the narrative summary below. Five papers on costs have been reported as contextual information.

The nine trials included in this review are the same studies included in the USPSTF review. For details on the individual trials see Characteristics of Included Studies (Appendix 6). Four studies were considered truly randomized, while five were quasi-randomized or provided incomplete information about randomization. GRADE tables were created for all nine studies and separately for the four truly randomized studies and the five others, which hereafter are treated as quasi-randomized studies. The Edinburgh trial was excluded from this review as there were inconsistencies with allocated practices getting the correct allocation and several practices changed allocation.

Results for Key Questions

Key Question 1a: Does screening with mammography (film and digital) or MRI decrease breast cancer mortality and all cause mortality for women aged 40–49 and ≥70?

The USPSTF review provided an updated analysis of the meta-analysis from the 2002 review with additional data from the AGE trial and Göteborg; both were assessed as fair in quality. The AGE trial (N=160,840) randomly assigned women aged 39 to 41 to an annual mammography screening until the age of 48, or to a control group that received usual care. Overall, 81% of the participants attended at least one screen; the mean number of screens was 4.5. At the follow-up of 10.7 years, the relative risk for breast cancer mortality among women assigned to screening was 0.83 (95% CI 0.66–1.04) and for all cause mortality the relative risk was 0.97 (95% CI 0.89–1.04). The Göteborg trial evaluated screening for women aged 39 to 59. Enrolled women (N=52,222) were randomly assigned to mammography screening every 18 months or to a control group that received usual care. Attendance at the first screening was 80%, and there was intention to treat analysis. Among women aged 39–49 at trial entry who had been randomized for screening, the relative risk of breast cancer mortality was 0.69 (95% CI 0.45–1.05) after 13 years of follow-up. The USPSTF 2009 meta-analysis resulted in a pooled relative risk for breast cancer mortality in

* The USPSTF review included eight studies; the Canadian National Breast Screening Study (CNBSS) was treated as a single trial. This review included nine studies because we separated the CNBSS into two trials: CNBSS-1 (40-49 years) and CNBSS-2 (50-59 years).
women assigned to mammography aged 39–49 of 0.85 (95% CI 0.75–0.96), which was consistent with the findings in the 2002 USPSTF review.4

Since the 2009 USPSTF review there have been no new screening trials. A Cochrane review on the effectiveness of mammography screening was published in 2009.18 The main objective of the Gotzsche and Nielsen review was to examine the effect of mammographic breast cancer screening on mortality and morbidity. Gotzsche and Nielsen identified 11 studies but excluded two because of small sample size68,69 and one because the intervention was a prevalence screening and there was biased randomization.70 Eight trials met the inclusion criteria and were included in the review; these are the same trials included in the USPSTF review. Three studies were considered to be adequately randomized,58-61 four were suboptimally randomized.63,64,66,71 One trial was inadequately randomized; data from this trial were excluded from the analysis.67

In the meta-analysis of the seven included studies from the Gotzsche and Nielsen review (adequate and suboptimally randomized) of women under 50 years, the pooled effect of screening versus no screening on breast cancer mortality was a relative risk reduction of 0.84 (95% CI 0.73–0.96) at the 13-year follow-up.18 Their meta-analysis of all seven trials was similar to the results of the USPSTF review which indicated a 15% reduction in breast cancer mortality in favor of screening (RR 0.85, 95% CI 0.75–0.96).3 However, when Gotzsche and Nielsen restricted their analysis to the three truly randomized studies, the point estimate for the reduction in breast cancer mortality due to screening was similar but was no longer statistically significant (RR 0.85, 95% CI 0.73–1.00).18 Gotzsche and Nielsen did not provide separate pooled data for women over the age of 70. The USPSTF review reported that for this age group there were insufficient data in the trials to perform a meta-analysis.3

In this 2010 meta-analysis (which included the same trials as the USPSTF and is summarized below in Table 2), screening led to a reduction in breast cancer mortality among women of all ages (RR 0.82, 95% CI 0.74–0.91) (Evidence Set 6) and women aged 39–49 (RR 0.85, 95% CI 0.75–0.96) (Evidence Set 1). The meta-analysis of the two trials that reported results for women aged ≥70 years (Swedish Two County, East and West) found that screening led to a nonsignificant reduction in breast cancer mortality (RR 0.68, 95% CI 0.45–1.01) (Evidence Set 2). This analysis is different than what is presented in the USPSTF review because our data are raw event rates whereas the USPSTF chose to use modeling data. We were unable to independently verify the event rates used for that analysis.

<table>
<thead>
<tr>
<th>Age, years</th>
<th>Truly randomized trials</th>
<th>RR for breast cancer mortality (95% CI)</th>
<th>GRADE quality rating</th>
<th>All trials (includes quasi-randomized)</th>
<th>RR for breast cancer mortality (95% CI)</th>
<th>GRADE quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>39–49</td>
<td>3</td>
<td>0.85 (0.73–1.00)</td>
<td>HIGH</td>
<td>8</td>
<td>0.85 (0.75–0.96)</td>
<td>MODERATE</td>
</tr>
<tr>
<td>50–59</td>
<td>2</td>
<td>1.00 (0.82–1.22)</td>
<td>HIGH</td>
<td>7</td>
<td>0.82 (0.68–0.98)</td>
<td>MODERATE</td>
</tr>
<tr>
<td>60–69</td>
<td>2</td>
<td>0.91 (0.74–1.11)</td>
<td>HIGH</td>
<td>5</td>
<td>0.69 (0.57–0.83)</td>
<td>MODERATE</td>
</tr>
<tr>
<td>70–74</td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
<td>0.68 (0.45–1.01)</td>
<td>LOW</td>
</tr>
<tr>
<td>All ages</td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td>0.82 (0.74–0.91)</td>
<td>MODERATE</td>
</tr>
</tbody>
</table>

Source: compiled from data presented in Evidence Sets 1 through 6.
Analysis of data for women aged 50 to 69 years showed screening led to a reduction in breast cancer mortality (RR 0.79, 95% CI 0.68–0.90) (Evidence Set 3). When meta-analysis was restricted to the two truly randomized studies, the reduction became statistically nonsignificant (RR 0.91, 95% CI 0.74–1.11) (Evidence Set 3). In strata defined by age, breast cancer mortality was reduced among women aged 50 to 59 years in all seven trials (RR 0.82, 95% CI 0.68–0.98), but not when the analysis was limited to the two truly randomized trials (RR 1.00, 95% CI 0.82–1.22) (Evidence Set 4).

For women aged 60–69 years, one trial was truly randomized and four were quasi-randomized. Combined in this meta-analysis these studies indicate a 31% reduction in breast cancer mortality (RR 0.69, 95% CI 0.57–0.83) (Evidence Set 5). For women aged 60–69, the USPSTF pooled data from two trials which resulted in a RR of 0.68 (95% CI 0.54–0.87). Gøtzsche and Nielsen reported a relative risk of 0.77 (95% CI 0.69–0.86) for all women screened in seven trials.

No RCT has assessed the effect on breast cancer mortality of screening with MRI for women of average-risk. Also there were no trials of digital or MRI screening.

Deaths ascribed to any cancer

The USPSTF review did not analyze the data for mortality ascribed to any cancer. Gøtzsche and Nielsen reported that the adequately randomized trials did not find an effect of mammography on any cancer deaths (RR 1.02, 95% CI 0.95–1.10) with a 10.5-year follow-up for the Canadian trials and a 9-year follow-up for Malmö. We found no new trials.

All cause mortality

The USPSTF review did not provide data on all cause mortality. Gøtzsche and Nielsen report that the trials had insufficient power to detect the effect of screening on all cause mortality. Pooled results from the first Canadian National Breast Screening Study (CNBSS-1) and the AGE trials showed that screening did not significantly reduce all cause mortality among a total of 211,270 women aged 39–49 (RR 0.97, 95% CI 0.91–1.04) (Evidence Set 7). The only other randomized trial that considered all cause mortality collected data on women aged 50–59 and reported that screening was associated with a relative risk of 1.06 (95% CI 0.96–1.18) (Evidence Set 7). Compared with analyses of breast cancer mortality, statistical power was limited for analyses examining all cause mortality.

Duration of follow-up

Each of the studies included in the review were long-term (more than 10 years), multi-follow-up trials. Women were randomized to receive mammography, on average on a 24-month basis, and it was repeated five to six times over the course of the trial. Table 3 and Table 4 present the results of mammography screening for women younger than 50 years and include mortality (both all cause and breast cancer) at two points in time. The mid-range follow-up represents a period of between six and eight years. The last published follow-up showing results for breast cancer and all cause mortality took place at a median of approximately 11 years.

Additional considerations

Gathering evidence for the effectiveness of mammography screening for women younger than 40 years or over the age of 75 was beyond the scope of this review. However, evidence for these age groups, if present in the literature, would have been located by our search. This search found no studies that met our inclusion criteria to support making recommendations for or against screening for these groups.
Table 3: Summary of Evidence – Mammography Trials for Women Younger than 50 Years – Breast Cancer Mortality

<table>
<thead>
<tr>
<th>Study</th>
<th>Age at Entry (Years)</th>
<th>Screening Interval (Months)</th>
<th>Year Study Began</th>
<th>Total Study Sample Size</th>
<th>Breast Cancer Mortality* at Mid-range Follow-up (6 to 8 Years)</th>
<th>Breast Cancer Mortality* at 14-Year Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Study Group</td>
<td>Control Group</td>
<td>Study Group</td>
</tr>
<tr>
<td>HIP^66</td>
<td>40–49</td>
<td>12</td>
<td>1963</td>
<td>13,740</td>
<td>13,740</td>
<td>19</td>
</tr>
<tr>
<td>Malmö^58</td>
<td>45–54</td>
<td>18–24</td>
<td>1976</td>
<td>13,568</td>
<td>12,279</td>
<td>N/A</td>
</tr>
<tr>
<td>Östergötland (E-County)^65</td>
<td>40–49</td>
<td>24</td>
<td>1977</td>
<td>10,285</td>
<td>10,459</td>
<td>15</td>
</tr>
<tr>
<td>Kopparberg (W-County)^55</td>
<td>40–49</td>
<td>24</td>
<td>1977</td>
<td>9,582</td>
<td>5,031</td>
<td>13</td>
</tr>
<tr>
<td>CNBSS-1^56</td>
<td>40–49</td>
<td>12</td>
<td>1980</td>
<td>25,214</td>
<td>25,216</td>
<td>38</td>
</tr>
<tr>
<td>Stockholm^64</td>
<td>40–49</td>
<td>28</td>
<td>1981</td>
<td>14,303</td>
<td>8,021</td>
<td>16</td>
</tr>
<tr>
<td>Göteborg^63</td>
<td>39–49</td>
<td>18</td>
<td>1982</td>
<td>11,724</td>
<td>14,217</td>
<td>16</td>
</tr>
<tr>
<td>UK AGE^61</td>
<td>39–41</td>
<td>12</td>
<td>1991</td>
<td>53,884</td>
<td>106,956</td>
<td>26</td>
</tr>
</tbody>
</table>

* From the 2002 USPSTF report Appendix 3, with the addition of updated numbers for Göteborg and the UK AGE study
<table>
<thead>
<tr>
<th>Study</th>
<th>Age at Entry (Years)</th>
<th>Screening Interval (Months)</th>
<th>Year Study Began</th>
<th>Total Study Sample Size</th>
<th>All Cause Mortality* at Mid-range Follow-up</th>
<th>All Cause Mortality* at 14-Year Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Study Group</td>
<td>Control Group</td>
<td>Study Group</td>
</tr>
<tr>
<td>CNBSS-159</td>
<td>40–49</td>
<td>12</td>
<td>1980</td>
<td>25,214</td>
<td>25,216</td>
<td>159</td>
</tr>
<tr>
<td>Göteborg</td>
<td>39–49</td>
<td>18</td>
<td>1982</td>
<td>11,724</td>
<td>14,217</td>
<td>178</td>
</tr>
<tr>
<td>UK AGE</td>
<td>39–41</td>
<td>12</td>
<td>1991</td>
<td>53,884</td>
<td>106,956</td>
<td>N/A</td>
</tr>
</tbody>
</table>

* From the 2002 USPSTF report Appendix 3, with the addition of updated numbers for CNBSS-1, Göteborg, and the UK AGE study
Key Question 1b: Does CBE screening decrease breast cancer mortality for women aged 40–49 and ≥70? Alone or with mammography?

The USPSTF concluded that the effectiveness of screening with CBE has not been established.3 This update did not identify any new studies of the impact of CBE (alone or with mammography) on breast cancer mortality.

Key Question 1c: Does BSE practice decrease breast cancer mortality for women aged ≥40?

The 2009 USPSTF3 review reported on the preliminary findings of two studies conducted in Russia and Shanghai.72,73 These trials reported that BSE did not lead to significant differences between BSE and control groups in breast cancer mortality or all cause mortality. Results from these studies (in women aged 39 years and older) were combined and showed little impact on breast cancer mortality (RR 0.98, 95% CI 0.83–1.15) (Evidence Set 8).

No new studies on the impact of BSE on breast cancer mortality were located in the updated literature search.

Key Question 2a: What are the harms associated with screening with mammography (film and digital) and MRI?

Harms from mammography screening include false-positives (discussed below) or false-negatives, overdiagnosis, unnecessary surgeries, radiation exposure, and psychological distress. False-negative results cause a delay in diagnosis for women who are subsequently found to have breast cancer.

The USPSTF review reported that published data on false-positive or false-negative mammography results were limited.3 The review reported that in women aged 40–49 there is a false-positive rate of up to 56% and a cumulative risk for all women of 21% to 49% after 10 mammography screenings. False-negative results are lowest and rates of additional imaging are most common in women aged 40–49 years.

False-positives

False-positive mammography results cause women who do not have cancer to be subjected to additional screening and needle or surgical biopsies.

The USPSTF review3 reported cumulative false-positive data from studies by Elmore et al.74 and Hofvind, Thorsen and Trelth.75 Elmore et al. used 10-year retrospective cohort data from 2,400 women who were between the ages of 40 and 69 years when they entered the study. A total of 9,762 screening mammograms were performed. The estimated cumulative risk of a false-positive result for all women after 10 mammograms was 49.1% (95% CI 40.3%–64.1%). Hofvind et al.75 used data from the Norwegian Breast Cancer Screening Program, in which all women aged 50–69 years are invited to biennial two-view mammography screening. False-positive estimates were based on the screening data for 83,416 women who participated in all three rounds of screening. It was estimated that women aged 50 to 51 who participate in three biennial screening rounds would have a 20.8% risk of a false-positive recall during a screening period of 20 years.
Using data from the Breast Cancer Surveillance Consortium, Hubbard, Miglioretti and Smith\(^48\) estimated the cumulative probability of a false-positive mammography screening result after the first, fifth, and tenth screening exams, for women aged between 40 and 59 years at the start of the study. The probability of false-positives at first mammography was 16.2% (95% CI 16.0%–16.4%) and was consistent across three different modeling techniques. By the fifth screening round, the models indicated a range of cumulative probability of false-positives from 40.7% (95% CI 40.3%–41.2%) to 52.8% (95% CI 52.5%–53.2%); and for the tenth screening round, a range of cumulative probability of 58.2% (95% CI 56.1%–60.4%) to 77.0% (95% CI 76.7%–77.3%), depending on the modeling strategy used.

Bluekens et al. examined referral rates in Dutch women when there was a transition from digital mammography to full-field digital mammography (FFDM).\(^76\) Their findings showed referral patterns peaked with the first period of FFDM when there was an 88% false-positive rate due to pseudo-lesions and increased detection of benign microcalcifications. There was a higher overall referral rate in FFDM screening in both the first and subsequent exams (\(p<.001\)).

Are women who receive false-positive results more likely to return for mammographic screening? There were 12 observational studies: two from Canada, five from Europe, and five from the United States included in a systematic review\(^77\) cited by the USPSTF. The quality of all 12 was judged to be very low in the GRADE rating. Individual studies were too heterogeneous to combine the effects for this question. However, the individual study reports found a range of results from more likely to less likely to return for subsequent screening; no conclusion can be drawn.

**Overdiagnosis**

Any invasive or noninvasive breast cancer detected by screening that would not have been identified clinically or would not have resulted in symptoms or death in a person’s lifetime is called overdiagnosis.\(^78\)

Determining levels of overdiagnosis are primarily based on data from randomized trials that have been abstracted and subjected to trend analysis or modeled data. Among the studies on overdiagnosis included in the USPSTF review, estimates of overdiagnosis for invasive cancer range from <1% to 30% in the screened population. Overdiagnosis of noninvasive cancer ranged from <1% to 37% for the screened population.\(^3\)

Our search located four primary studies and one systematic review that examined the question of overdiagnosis in breast cancer screening. Jørgensen, Zahl and Gøtzsche collected data on incidence of carcinoma in situ and invasive breast cancer in Danish women in all areas with and without screening over 13 years (1991–2003) and for the 20-year period prior to screening being introduced (1971–1990).\(^46\) For women aged 50–69 they reported a 35% rate of overdiagnosis when comparing unadjusted incidences for the screened and nonscreened areas. The adjusted Poisson regression analysis indicated a relative risk of 1.40 (95% CI 1.35–1.45) for the entire screening period. There was a potential compensatory drop in women aged 70–79 (RR 0.09, 95% CI 0.88–0.96); the study authors suggest that the most reliable estimate of overdiagnosis is 33%.

Duffy et al.\(^47\) estimated the number of breast cancer deaths prevented and the rate of overdiagnosis in mammography screening programs for women aged 50–69 by re-examining data from the Swedish Two-County Trial and the UK Breast Screening Programme (UKBSP) in England. Their estimates of absolute benefits of screening over 20 years were 8.8 (Swedish Two-County Trial) and
5.7 (UKBSP) breast cancer deaths prevented for every 1,000 women screened. The corresponding overdiagnosis rates were 4.3 and 2.3 per 1,000 over 20 years.

Morrell et al. estimated overdiagnosis of invasive breast cancer in screening programs in New South Wales, Australia. This study examined incidences and trends of invasive breast cancer in both screened and unscreened populations and compared expected incidence in 1999–2001 with observed incidence for the same period to calculate overdiagnosis. Linear regression modeling was used to estimate invasive breast cancer for women without screening. This study estimated overdiagnosis among women aged 50–69 years in New South Wales to be 42% and 30% using interpolation and extrapolation methods, respectively.

A study in Catalonia, Spain, modeled incidence of invasive breast cancer and overdiagnosis for a cohort born between 1935 and 1955. Their estimate of overdiagnosis ranged from 0.4% for women born in 1935 to 46.6% for women born in 1950.

A systematic review examined secular trends in breast cancer incidence and overdiagnosis. In the absence of clinical trials with a lifelong follow-up, Jørgensen and Gøtzsche reviewed the literature to identify trends in the incidence of breast cancer before and after mammography screening to estimate the extent of overdiagnosis. They searched PubMed and identified five studies with relevant information. Data were presented from the United Kingdom; Manitoba, Canada; New South Wales, Australia; Sweden; and parts of Norway. Their results indicated that in populations offered organized breast cancer screening, overdiagnosis (including that of carcinoma in situ) was 52% (RR 1.52, 95% CI 1.46–1.58). Overdiagnosis in publicly organized mammography screening programs could not be calculated from the systematic review of incidence trends because the study did not provide data with which to estimate expected annual incidence of breast cancer.

**Unnecessary Biopsies or Surgeries**

Table 5 presents the estimated number of Canadian women with benign findings on surgical or percutaneous breast biopsy performed as follow-up to screening mammography.

**Table 5: Estimated Number of Women with Adverse Outcomes Following Screening Mammography**

<table>
<thead>
<tr>
<th></th>
<th>40 – 49 y</th>
<th>50 – 69 y</th>
<th>70 – 74 y</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Per 1,000 women screened</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False positive mammograms</td>
<td>327</td>
<td>282</td>
<td>212</td>
</tr>
<tr>
<td>Unnecessary biopsies*</td>
<td>36</td>
<td>37</td>
<td>26</td>
</tr>
<tr>
<td><strong>Per one death prevented</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number needed to screen</td>
<td>2,108</td>
<td>721</td>
<td>451</td>
</tr>
<tr>
<td>False positive mammograms</td>
<td>690</td>
<td>204</td>
<td>96</td>
</tr>
<tr>
<td>Unnecessary biopsies*</td>
<td>75</td>
<td>26</td>
<td>11</td>
</tr>
</tbody>
</table>

Notes: Results are expressed per thousand women screened for a median of 11 years (estimated as a total of 4 screening mammograms per woman assuming a screening interval of 2-3 years). The duration of 11 years was chosen because it was the approximate median duration of follow-up during the included randomized trials. Data assume that rescreen rates stay constant over time. Some data that were used in these calculations were not available for Alberta. Cancer detection rates which were used in these calculations may vary in provinces where screening frequencies differ.

* percutaneous or surgical biopsies done in a woman subsequently found not to have cancer
In the Gøtzsche and Nielsen review, for the three truly randomized trials, the relative risk of mastectomies and lumpectomies in the mammography screening group was 1.31 (95% CI 1.22–1.42), clearly indicating that those screened had more procedures (Evidence Set 9). This was similar to the results that also included the two quasi-randomized trials (RR 1.42, 95% CI 1.26–1.61) (Evidence Set 9). Gøtzsche and Nielsen found that the increased surgery rate could not be explained simply by the detection of tumors. The argument that screening allows less aggressive treatment (lumpectomies versus mastectomies) was not borne out, in that the relative risk of mastectomy alone was 1.20 (95% CI 1.08–1.32) among those in the mammography group (Evidence Set 9).

**Radiation Exposure**

The USPSTF reported that no trials directly measured the association between mammography and radiation exposure.

A 2010 paper examined radiation dose and cancer risk from breast imaging studies. This paper used the mean glandular dose (MGD) for two-view single-film mammography (SFM) and digital mammography (DM) from peer-reviewed literature (the search strategy was not described) to estimate the average MGD and the range of MGDs to the US screening population. Two-view DM and SFM involve average MGD radiation doses of 3.7 and 4.7 mGy, respectively. These are associated with a lifetime average risk (LAR) of fatal breast cancer of 1.3 and 1.7 cases per 100,000 women aged 40 years at exposure and less than one case per million women aged 80 years at exposure. Annual screening digital or screen-film mammography performed in women aged 40–80 years is associated with a LAR of fatal breast cancer of 20 to 25 cases in 100,000.

**Anxiety, distress, and other psychological responses**

We retrieved two reviews cited by the Nelson review that were related to anxiety, distress, and other psychological responses. The Brett review had no new data and not enough information to create GRADE tables. As stated by Brett, “Studies used a range of measures, the measures were not used in a uniform way, and different time intervals were used.” Most studies had no comparison groups. Brett comments on the papers assessing other psychological impact indicators: “due to the heterogeneous nature of these outcomes, they have not been tabulated but have been included in relevant results sections.” The conclusions in this paper state that “mammographic screening does not appear to have a negative psychological impact for the majority of women who receive an initial clear result after screening. However, for women who are recalled for further investigations after screening there are significant adverse psychological consequences in the short term, which may remain to a lesser extent long-term.”

A 2010 meta-analysis examined the effect of false-positive mammography on generic and breast cancer-specific psychosocial outcomes of women (distress about breast cancer, somatization or symptoms in the breast, fear of getting breast cancer, anxiety about breast cancer, worry about breast cancer, perceived likelihood of breast cancer, perceived benefits of mammography, frequency of BSE). This meta-analysis included 21 papers representing 17 studies published between 1989 and 2007. The study samples contained usable data for 20,781 participants (study sample range 89 to 9,578). Data were pooled to determine effect size for psychological effects of false-positive mammograms (Table 6).
Table 6: Psychological Effects of False-Positive Mammograms

<table>
<thead>
<tr>
<th>Effect</th>
<th>Increase Effect Size (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distress</td>
<td>0.16 (0.10–0.22)</td>
</tr>
<tr>
<td>Somatisation</td>
<td>0.12 (0.05–0.19)</td>
</tr>
<tr>
<td>Fear</td>
<td>0.08 (0.03–0.14)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.22 (0.18–0.27)</td>
</tr>
<tr>
<td>Worry</td>
<td>0.12 (0.08–0.16)</td>
</tr>
<tr>
<td>Perceived likelihood of getting breast cancer</td>
<td>0.09 (0.04–0.14)</td>
</tr>
<tr>
<td>Perceived benefits of mammography</td>
<td>0.11 (0.06–0.17)</td>
</tr>
<tr>
<td>Frequency of BSE</td>
<td>0.11 (0.04–0.19)</td>
</tr>
</tbody>
</table>

Key Question 2b: What are the harms associated with CBE?

Harms of CBE can include false-positive results that can lead to further imaging. As well, considerable anxiety and distress are associated with false-positives. False-negative results from CBE can lead to delay in a cancer diagnosis.

The USPSTF’s review reporting on harms associated with CBE was limited to three studies, a pilot study, a study that ended early because of low participation, and one case control study. The case control study showed that of the 485 women who received CBE within one year prior to a breast cancer diagnosis and within 15 years of death attributed to breast cancer, CBE failed to detect breast cancer in four out of five cases.

Our search located one additional study that examined the harms of CBE. In a cohort study of women being screened, 232,515 participants received CBE and mammography and 57,715 participants received mammography alone. CBE was offered as well as mammography in nine regional cancer centres and at 59 affiliated centres. In those centres, the cancer detection rate for mammography referrals was 5.9 per 1,000, while for CBE and/or mammography the detection rate was 6.3 per 1,000 referrals. The false-positive rate for mammography referrals was 6.5%; for CBE and/or mammography referrals the false-positive rate was 8.7%. With CBE, an additional 0.4 cancers were detected per 1,000 women screened relative to mammography alone, while there was a 2.2 percentage-point increase in the false-positive rate. In other words, for every 10,000 women screened, there would be an additional four cancers detected and of the 9,937 women without cancer, there would be an additional 219 false-positives. For each additional cancer detected with CBE per 10,000 women, there would be 55 additional false-positives.

Key Question 2c: What are the harms associated with BSE?

Harms associated with BSE are similar to those outlined in the section above on CBE.

The USPSTF’s review reported trials in Russia and Shanghai that found women assigned to BSE had a higher incidence of benign biopsy results than women in the control group: RR 2.05 (95% CI 1.80–2.33) in the Russian trial and RR 1.57 (95% CI 1.48–1.68) in the Shanghai trial.
Our literature search located no new studies that examined the harms associated with BSE as a screening method for breast cancer.

**Results for Contextual Questions**

Five new reports of costs were found as well as three systematic reviews and 23 primary studies related to patient preferences and values.

**Contextual Question 1: What is the cost-effectiveness of screening?**

Five new studies related to cost-effectiveness were found in this update; all were reports of micro-simulation modeling.\(^{51-55}\)

Ahern and Shen assessed the cost-effectiveness of screening schedules recommended by the three major US cancer organizations and compared them with alternative strategies.\(^{52}\) Costs of screening examinations, subsequent work-up, biopsy, and treatment after diagnosis were all considered. Incremental cost-effectiveness ratios were used to compare strategies. Mammography and CBE alternating years from ages 40 to 79 was a cost-effective alternative compared to guidelines of the National Cancer Institute (mammography every one to two years), American Cancer Society (annual mammography for women 40 years and older, CBE every three years beginning at age 20, annually at age 40) and USPSTF (mammography every one to two years, with or without CBE for women 40 and older) and cost an additional USD 35,500 per quality-adjusted life year (QALY) saved compared with no screening. The American Cancer Society guideline was the most effective and the most expensive, costing an additional USD 680,000 for an added QALY compared with the above alternative.\(^{52}\)

Another modeling study considered cost-effectiveness of opportunistic versus organized mammography screening for women aged 50 to 69 years in Switzerland.\(^{53}\) Assuming an 80% participation rate and compared to no screening, both yielded a similar reduction in breast cancer mortality (13%) during the lifespan of the population screened and a similar reduction in predicted breast cancer mortality rate (25%) 20 years after the start of the program. The 3% discounted cost-effectiveness ratio for organized screening was €11,512 per life year gained while opportunistic screening had twice the cost, with a ratio of €22,671 to €24,707 per life year gained.\(^{53}\)

The peak incidence for breast cancer in women in the Republic of Korea occurs between ages 45 and 49. A micro-simulation modeling exercise was done to determine the most cost-effective screening interval and target age range for Korean women, from the perspective of their national healthcare system. The most cost-effective strategies were: biennial mammography screening for women at least 40 years of age; and biennial screening beginning at age 35.\(^{57}\) It is not known to what extent this finding would apply to women of Korean descent living in Canada.

Two recent papers compared film and digital mammography. Wang et al.\(^{54}\) compared costs within the Australian healthcare system, using 2007 prices for Australian dollars. They concluded that there is no evidence that digital and film-screen differ significantly in terms of diagnostic accuracy on a population level in either screening or diagnostic use, so cost comparison alone was done and not cost-effectiveness analysis. They found that digital mammography cost $11.40 (Australian) more per examination in the screening setting compared to film-screen mammography. They did further analysis of cost-effectiveness of digital mammography for diagnosis. The other comparison was done by the Canadian Agency
for Drugs and Technologies in Health (CADTH),\(^{55}\) which set out to review the cost-effectiveness of digital versus film mammography. CADTH concluded that in terms of clinical effectiveness, full field digital mammography and film mammography appear equivalent. However, they were unable to draw conclusions about the relative costs of the technologies; costs varied across studies as they utilized different time horizons.

**Contextual Question 2: What are patient preferences and values with regard to breast cancer screening?**

Three systematic reviews and 23 primary studies (surveys, time trade-off studies, and qualitative studies) were identified as relevant to this review of patient preferences and values. Preferences regarding aspects of screening, patient-physician involvement in decision making, and other factors related to screening or intention to be screened are included in this section. Some of the 23 primary studies were included in the systematic reviews and most will not be addressed individually.

**Patient preferences for breast cancer screening**

A systematic review of preferences for cancer screening found eight studies; these were related to breast cancer (three studies), colorectal cancer (four studies), or both (one study).\(^{84}\) Most were based on contingent valuation or willingness to pay and involved the general public. Participants valued test accuracy and mortality reduction and did not consider potential harms of testing. Of particular interest, a study included in this review was a random sample of 207 Danish women >50 years, interviewed to determine preferences of screening program characteristics.\(^{85}\) Each participant was presented with consequences of no screening versus three alternative screening programs in terms of numbers of mammography exams over the next 25 years, risk of dying from breast cancer in the next 25 years, risk of calls for further unnecessary exams, and out-of-pocket expenses and the consequences of no screening. In the discrete ranking analysis, significant results were found for expectation of risk reduction and participant’s education level, while risks of false-positives and out-of-pocket expenses had an impact of decreasing preferences for screening.

Two groups of women, those with BRCA and controls, from two centres (one in the United States and one in Toronto) indicated how many years of life expectancy they would trade to avoid BRCA mutations, breast/ovarian cancer, and five preventive measures including prophylactic surgery, annual mammograms, and annual MRI.\(^{86}\) Both groups of women gave mammography and MRI the highest trade-off values (most favourable), considering them to have little impact on the quality of their lives. Standard deviations of ratings were high, indicating the variation in individual preferences and the need to consult with individual women in treatment decisions.

In a survey of 1,528 US women at the time of a screening appointment, 97% believed that a false-positive result would not deter them from continuing with regular screening.\(^{87}\) Most would have been willing to be recalled more often for either a noninvasive (86%) or an invasive (82%) procedure if it might increase the chance of detecting a cancer earlier. Women under 60 years and those previously recalled were more willing to be called back more often for a noninvasive or an invasive procedure. Women preferred the inconvenience and anxiety associated with a higher recall in return for a possibility of detecting breast cancer earlier.\(^{87}\) Another survey of US women, 41 to 70 years and attending mammography screening, indicated willingness to undergo
mammography even if the benefit was reduced; about half would not have mammography if no clear benefit existed, and about 24% would still have it if it increased the chances that the breast could be preserved.88

**Patient-physician involvement in decision making**

Most guidelines recommend that healthcare providers discuss the indicators for and against screening and make the decisions with individual patients. In a younger group, US women aged 40 to 44 who presented for screening preferred to make the screening decision after considering their medical provider’s opinion (38%) or together with their medical providers (46%); fewer than 10% preferred that the decision be made by the woman or her provider alone.89 Women aged 50 to 69 years residing in Geneva considered that the decision to undergo mammography screening should be made by the doctor alone (5.6%), doctor primarily (42.6%), woman and doctor sharing equally (45.0%), woman primarily (4.2%), or woman alone (2.4%).90

A US survey investigated nine common medical decisions men and women face (including breast cancer screening for women and prostate cancer screening for men).91,92 Respondents made more decisions in the past two years if they had a primary care provider or poorer health status; and fewer decisions if they had lower education, were male, or were under 50 years. Most patients reported that providers made a recommendation, and that those recommendations generally favoured taking medical action. Forty percent of women reported that their providers asked them about their preferences for breast cancer screening, and 20% reported that providers discussed reasons for not having breast screening. Patient confidence in decisions was higher among patients who had made the decisions themselves or who had been asked their preference, and lower when patient-provider discussions had included cons of screening.91

Women in New South Wales (94.6%) preferred to share mammography decision making equally with their doctor or to take a more active role, with only 5.4% reporting they wanted the doctor to make these decisions on their behalf.93 While this pattern was consistent across all age groups, women who had a usual doctor were more likely to report having an active role in decision making. Most women wanted information about the possibility of false test results (91.5%) and test side-effects (95.6%), stating that this information would make them anxious, but they wanted the information anyway. However, about a third of the women reported that their doctor never provided this information.

A qualitative study of women over 80 years and their physicians found that women were divided between enthusiastic, opposed, and undecided about continued mammography screening.94 The undecided were most influenced by physician recommendations, but the physicians were uncomfortable having discussions with these patients about stopping screening.

Women aged 70 years and older who had regularly participated in mammography screening in New South Wales, Australia, were eligible to participate in a trial of the effectiveness of a decision aid about whether to continue or stop mammography screening.95 Women received a decision aid that provided balanced, quantitative information or standard information available from the screening program. Women who received the decision aid had an increased knowledge score and slightly reduced decisional conflict with no increase in anxiety and no change in participation in screening compared to controls.
Other factors influencing the decision or intention to be screened

Participation rates in screening vary by geographic location, by income and education levels, by ethnicity, by satisfaction with previous screening examinations, and by a variety of other factors. Ackerson and Preston conducted a review of studies of women’s decisions to have breast and cervical screening.96 They performed content analysis on the 19 papers and found three recurrent themes:

1. Fear: depending on the source of the fear, women were shown to avoid (when fearing the test or the results) or to seek (when fearing cancer itself) screening; in both cases, they acted to reduce their fear.

2. “I take good care, I can detect cancer in my body”: many women think that routine screening is unnecessary because they take good care of themselves and do not experience symptoms.

3. “No one told me that I should”: reflected by women who said they would attend screening if their care provider recommended it.96

This review included a study of women in Toronto, aged 25 to 45 years, and their intention to use mammography.97 Intention was not related to knowledge or availability of services but was related to care provider recommendation. Intention not to use was related to fear of radiation exposure, other daily duties taking priority, and belief in faith/destiny.

The themes of the review were supported in a single study of African American and White American women, aged 40 to 79 years, who had one mammogram.98 Across race, age, and family breast cancer history, women who believed that they were “very likely” to develop breast cancer were less likely to be re-screened than women who believed that their susceptibility was moderate (adjusted OR 2.83, 95% CI 1.51–5.30), and the effect was stronger in older women. Women aged 40 to 49 years (but not those aged 50 to 79 years) who believed they were “not likely” or “a little likely” to develop breast cancer were also less likely to be re-screened than those who reported moderate susceptibility.

For Tamil women from Sri Lanka, living in Toronto, the most common barriers to screening reported by the women were: lack of understanding of the role of early detection in medical care, religious beliefs, and fear of social stigmatization.99 Vietnamese Canadians have low screening participation rates for breast cancer; they reported barriers such as embarrassment and privacy of breasts (cultural influence that no other person should touch their breasts); if they are healthy enough to work, they do not need the examination; and illness as destiny that cannot be changed.100

A qualitative study of women in Sweden found six main themes that were important issues in reasoning about attendance or non-attendance at mammography screening: negative experiences; perceived risk factors; knowledge of one’s own body; perceived problems with mammography; political, ideological, and moral reasoning; and involuntary non-attendance due to the inability of the screening program to cover some women.101

Other important factors have been identified related to mammography uptake and re-screening. Ease of appointment scheduling (fitting with their time availability) is related to greater uptake,102 while factors related to reduced screening or re-screening are overall psychological distress,103 being younger (40–49),104 rating health as poorer and having lower satisfaction with previous
mammography experience,\textsuperscript{104} a desire for a holistic screening approach that does not separate breast from the rest of the body (a qualitative study of African American women),\textsuperscript{105} and feeling healthy so having no reason to be screened (Hispanic women in the United States).\textsuperscript{106}

Asian American women (women of Chinese, Korean, Filipino, and Asian Indian origin) have lower incidence of and mortality from breast cancer, and low uptake of mammography, CBE, and BSE.\textsuperscript{107} However, they are more likely to be diagnosed at an advanced stage. Wu et al. found 23 studies that identified several demographic variables that were consistently associated with mammography: insurance status, recency of physical examination, physician’s recommendation, and length of US residency.\textsuperscript{107} These demographic variables also were shown to be correlated with CBE.

Peipins and colleagues assessed the satisfaction of US women aged 19 and older with five screening examinations, including CBE and mammography.\textsuperscript{108} Women were very satisfied with care received during all screening exams. Women were more satisfied during CBE and physical exams if they perceived these exams as informative, clear, and complete and if they perceived the providers as informative and responsive to them when they asked questions. Women also were more satisfied with all three exams when the providers were perceived as relaxed during these exams.

The only study related to BSE was a qualitative study of Canadian women that identified factors related to conduct of BSE.\textsuperscript{109} Reluctance to perform the exam was influenced by participants’ perceptions of breast cancer as a lethal disease, the perceived threat it posed to their femininity, and their ability to negotiate an increasingly medical and technological healthcare system. Regarding the latter point, the women identified that they avoided doing BSE because they could not then be held personally accountable for the disease identification or progress.

In summary, most women value mammography in particular for perceived reduction of mortality; few women consider issues of further testing or harm arising from false-positives in their decision making. However, many of the studies were done when participants were already in screening programs. Other women refuse breast cancer screening because of fear, fatalistic beliefs, absence of symptoms, or work or family responsibilities that do not allow for daytime appointments. The majority of women prefer to be jointly involved in decision making with their care providers, but some would go for screening if recommended by their providers.

**Contextual Question 3: What is the effectiveness of screening for specific subpopulations?**

The CTFPHC has an interest in exploring breast cancer screening experiences of specific populations within the Canadian context. With this in mind, three groups were identified as having a unique Canadian perspective: Aboriginal women, women who reside in rural and remote locations, and women who are immigrants to Canada. The focused search for systematic reviews and key grey literature located seven papers that provided information for these identified populations. There are few data about cancer screening among First Nations or any ethnic groups in Canada because cancer registries in Canada do not routinely gather information about ethnicity.\textsuperscript{110,111}
Aboriginal Women

Burden of disease
Cancer is a leading cause of mortality in First Nations people, ranking third after heart disease and accidents, suicides, or homicides.112 Rates* of breast cancer in First Nations** women are lower than the rates for the general population of non-Aboriginal women; however, the incidence is rising.110,111,113 Breast cancer mortality is lower for First Nations women (11 per 100,000 compared with 25 per 100,000 for the general population), but that is because the incidence is lower (the incidence and mortality rates are rising at the same rate as in the general population).111 In Saskatchewan, for instance, incidence of breast cancer in Aboriginal women has been lower than in the general population but is now the same as in the general population.111 It is less likely that breast cancer in a First Nations woman will be diagnosed through screening, and First Nations women die more quickly following a diagnosis of breast cancer. Potential reasons for these poorer outcomes include lack of access to healthcare, cancer diagnosis at a later stage, higher level of co-morbidity, genetics, and/or lifestyle.110,114

Barriers to screening
Several reasons have been suggested for why Aboriginal women do not get breast cancer screening.110 Like many non-Aboriginal women, First Nations women reported thinking that screening was not necessary. Part of this attitude can be attributed to the lower incidence of breast cancer in First Nations women. Transportation and logistical deterrents exist for many Aboriginal women, especially those living in rural and remote reserves. Mammography is offered in regional cancer centres, large health facilities, and primary care. While mobile units exist, the sensitivity of the machines is such that they can be transported only on paved roads (only Québec has a flying mobile unit). Women who live where there are dirt roads have to travel to the nearest mobile unit or regional centre mostly at their own expense. Lack of access to consistent primary healthcare providers is also problematic. However, in one New Brunswick study it was demonstrated that when there is a family doctor (even working part-time) who recommended screening, the rate of screening for First Nations women was equivalent to that for the non-Aboriginal women in their community.115 This is similar for all Canadian women. A cross-sectional survey of 15,195 women aged 50 to 69 years found that women with a family doctor were twice as likely to have a mammogram as those without a family doctor.116

Rural and Remote-Dwelling Women
Statistics Canada reported that in 2008, 66% of women resided in urban areas while 34% resided in rural locations.26 Overall, rural women had mammography screening at a slightly lower rate than their urban counterparts (71% versus 73%).26 The difference between rural and urban usage of mammography screening varies significantly between provinces. The use of mobile units and educational and awareness campaigns has led to an increase in mammography screening for women in British Columbia, Manitoba, Saskatchewan, and New Brunswick.117

* National data were not available in retrieved literature.
** Separate data for Métis and Inuit were not present in the retrieved literature.
Barriers to screening

Healthcare and preventive programs such as screening can be difficult for women residing in rural and remote areas to access. Mobile clinics are available and have been successful at increasing the availability of the service; however, transportation issues may make access to mobile clinics difficult. All women residing in remote areas may face similar barriers as those described previously for Aboriginal women. In the Far North, some services are not available at all. In Nunavut, where there is no organized mammography program, only 32% of women reported having a mammogram in the last two years. Transportation is complicated in the North, where people often have to fly to access specialized services such as mammography. Not only is the cost prohibitive but many flights are cancelled because of snow and ice. Northern women have reported negative experiences with healthcare providers. While this could be true regardless of geographic location, women living in remote areas have less access to alternatives; therefore, they simply drop out rather than deal with a system they see as hostile.

Ethnicity

No Canadian cancer registry databases currently gather cancer incidence or mortality rates by ethnicity. This section is limited by the available literature. One of the reviews identified for patient preferences and values (Contextual Question 2) reported that Asian American women (women of Chinese, Korean, Filipino, and Asian Indian origin) have lower incidence of and mortality from breast cancer, and low uptake of mammography, CBE, and BSE, but are more likely to be diagnosed at an advanced stage. However, a more recent study suggests that breast cancer incidences for US-born Chinese, Japanese, and Filipina women are approaching those of non-Hispanic White women. Moreover, the incidence of breast cancer is continuing to increase among Asians living in their native countries. Survival after breast cancer diagnosis is shorter among foreign-born than US-born Asians.

Barriers to screening

It was reported that not speaking French or English is the most common reason for women not to participate in screening. Immigrant women may hold personal beliefs and cultural practices that do not lend themselves to preventive services like mammography screening. We refer the reader back to the preferences and values section of this report for more detailed information about beliefs and screening.

Contextual Question 4: What is the evidence of optimal frequency of screening with mammography?

Optimal mammography screening intervals

In the trials included in the USPSTF report the screening intervals varied from annual to a maximum of 33 months. The CNBSS-1 and -2 screened annually for five years; the HIP trial screened women annually for three years, and the AGE study screened women annually for a maximum of six years. The Stockholm trial screened women at intervals of 24 or 28 months. The Swedish Two-County Trial screened women aged 40–49 on average every 24 months and women 50–69 years every 33 months (Table 7). The Malmö trial screening intervals were 18 to 24 months for five rounds. Our search did not locate any analysis of the impact on mortality based on screening frequency between or among these studies. Pooled analyses examining the effect of screening on breast cancer mortality (stratified by screening frequency) are presented below. There was no statistical evidence that the benefit of screening differed for intervals of ≥24 months compared with intervals of <24 months.
### Table 7: Relative Risk of Breast Cancer Mortality for Mammography Screening Intervals

<table>
<thead>
<tr>
<th>Age, Years</th>
<th>(Screening Interval ≥24 Months)</th>
<th>(Screening Interval &lt;24 Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trials Included</td>
<td>RR for Breast Cancer Mortality (95% CI)</td>
</tr>
<tr>
<td>39–49</td>
<td>b, e, f</td>
<td>1.04 (0.72–1.50)</td>
</tr>
<tr>
<td>50–69</td>
<td>b, e, f</td>
<td>0.67 (0.51–0.88)</td>
</tr>
<tr>
<td>≥70</td>
<td>b, f</td>
<td>0.68 (0.45–1.01)</td>
</tr>
<tr>
<td>All Ages</td>
<td>b, e, f</td>
<td>0.77 (0.58–1.03)</td>
</tr>
</tbody>
</table>

** no trials for screening interval < 24 months at age 70+

Source: based on data presented in Evidence Set 10

Further stratified analyses suggested that the benefit of screening appeared similar in trials with screening intervals of 33 months [two trials62,71 with 98,431 women (RR 0.70, 95% CI 0.45–1.09)], with screening intervals of 24 months or greater [three trials62,71,71 with 193,905 women (RR 0.77, 95% CI 0.58–1.03)], and with annual screening [four trials56,57,61,66 with 311,165 women (RR 0.87, 95% CI 0.77–0.99)]. The small number of women screened in the 33-month group did not permit further stratification by age.

We searched Medline from 2000 forward to locate systematic reviews or primary studies examining the question of optimal screening frequency. From that search, one randomized controlled trial was located. The United Kingdom Co-ordinating Committee on Cancer Research (UKCCCR) directly compared different screening intervals.122 Women aged 50 to 62 (N=76,022) who attended a prevalent screen were allocated to the study arm (n=37,530) and were invited to three additional annual screens or were allocated to the control arm (n=38,492) and received the standard screen three years later. The endpoint was the predicted deaths from breast cancer. The prediction was based on the Nottingham Prognostic Index (NPI) and a method developed from the survival data in the Swedish Two-County Trial (2CS). The risk of death from breast cancer for the annual group was not significantly different from that for the three year group (RR 0.95, 95% CI 0.83–1.07 using the NPI; and RR 0.89, 95% CI 0.77–1.03 using the 2CS). The total predicted deaths were 36% in the study arm and 38% in the control arm.
Discussion

In the interval between the publication of the 2009 USPSTF report and October 2010, there have been no new trials to contribute to the debate about the effectiveness of breast cancer screening by mammography, CBE, or BSE. The USPSTF and Nordic reviews created considerable controversy, and the same trials included in those reviews have been analyzed in various ways since. Many of the included trials have been criticized for methodological issues such as inadequate randomization and inconsistent reporting of trial denominators. The Gotzsche and Nielsen review provide a good exploration of these issues.

This current review presents sensitivity analysis with all trials, with only the truly randomized trials, or with only the quasi-randomized trials included. In pooled analyses of 348,219 women aged 39–49, mammography reduced the risk of breast cancer mortality (RR 0.85, 95% CI 0.75–0.96; I²=0%). In women ≥70 years (n=17,646), pooled analyses showed a borderline but nonsignificant reduction in breast cancer mortality (RR 0.68, 95% CI 0.45–1.01; I²=0%). Meta-analysis of 250,274 women aged 50–69 confirmed a significant reduction in breast cancer mortality (RR 0.79, 95% CI 0.68–0.90; I²=41%). The overlapping confidence intervals for the age-specific estimates of the effect of mammography on breast cancer mortality preclude an assessment of whether mammography is truly more or less effective in younger or older women. The most reliable estimate of the true benefit of mammography may therefore be the pooled relative risk from all ages combined. Combined analysis of women (n=616,757) in the nine trials found a significant reduction in breast cancer mortality (RR 0.82, 95% CI 0.74–0.91; I²=36%). The available evidence did not permit assessment of whether screening mammography reduced all cause mortality in any age group.

Although mammography does appear to reduce the relative risk of breast cancer mortality, the absolute benefit of mammography is small. For example, the absolute number of deaths prevented per million women screened ranged from 474 among women aged 40–49 to 1,387 among those aged 50–69. The optimal frequency of screening cannot be determined at present but available data suggest no significant difference between screening intervals of one year and three years.

Gathering evidence for the effectiveness of mammography screening for women younger than 40 or over the age of 75 was beyond the scope of this review. However evidence for these age groups, if present in the literature, would have been located by our search. This search found no studies that met our inclusion criteria to support making recommendations for or against screening for these groups.

No evidence was found to support the benefit of CBE or BSE – either alone or in conjunction with mammography.

Harms for mammography screening are important and significant. The USPSTF reported women aged 40–49 have a false-positive rate of up to 56% and cumulative risk for all women ranging from 21% to 49% after 10 mammographic screens. The estimated cumulative probability of false-positives found in the new literature range from 58% to 77% after 10 mammographic screens. False-positive rates for CBE and BSE are higher than for mammography. In populations offered mammography screening, the relative risk of unnecessary surgeries (mastectomies and lumpectomies) was 1.31 (95% CI 1.22–1.42). It is estimated that Canadian women who undergo four screening mammograms will have false-positive rates of 330 (40–49 years), 280 (50–69 years), and 210 (70–74 years) (all rates expressed per thousand women screened every two to three years for a median of 11 years). During the same time and with the same screening intervals, 36 (40–49 years), 37 (50–69...
years), and 26 (70–74 years) benign percutaneous core biopsies will be reported per thousand women.79 Mammography screening does not seem to have adverse psychological consequences except in those recalled for further investigation; long-term effects are not clear. However, women who have experienced a false-positive reading will have higher levels of anxiety and fear related to the possibility of having a breast cancer diagnosis. Further, screening itself produces some risks: annual screening with digital or screen-film mammography performed in women aged 40–80 is associated with a lifetime average risk of fatal breast cancer of 20 to 25 cases in 100,000.44

Cost-effectiveness studies vary in their years of data collection and currency used. They focus on different aspects and use different modeling techniques. It is difficult to form firm conclusions based on available studies.

Qualitative studies have found that women value mammography for perceived reduction in mortality. Few women consider issues of further investigation or harms of screening. The majority of women want to be jointly involved with their healthcare providers in making the screening decision.

In the Canadian context, screening for breast cancer is lower in women who are Aboriginal, reside in rural and remote locations, or are new immigrants to Canada. Many have reduced access to mammography and are without family doctors, and some hold cultural beliefs that make it less likely that they will participate in breast cancer screening.

Digital mammography is now being implemented in most Canadian radiology departments, and treatment changes continue annually with dramatic changes in the last five years (e.g., targeted therapies, aromatase inhibitors, and partial breast radiation). Overdiagnosis is difficult to document in population studies, and the biology of breast cancer is difficult to ascertain from the existing screening trials. Moreover, breast cancer in younger women usually is more aggressive and often estrogen receptor–negative and has a shorter sojourn or lead time. Therefore biennial and even annual screening may miss interval cancers.

The aging Canadian population further complicates screening recommendations. The projections show that population aging, which has already begun, will accelerate in 2011 when the first baby-boom cohort (born in 1946) reaches the age of 65. This rapid aging is projected to last until 2031, when seniors will account for 23% to 25% of the total population. This would be almost double their current proportion of 13%.8 Population projections indicate a significant increase in women over age 70, and with increased life expectancy this also impacts screening and treatment of elderly women. This may lead to epidemiologic transition and more effective screening of women in the 60 to 69 year cohort. However, the impact on women over age 70 is less clear.123

**Limitations**

There are several limitations associated with this review. First, the search was limited to only those databases searched in the USPSTF review; only English language papers were included in the USPSTF search, and only English and French were included in this update; only Medline and Cochrane databases were searched. EMBASE would be a logical database for searching for this question, but this was not done for the current review.

Second, the searches for information about patient preferences and values and about special populations were focused and limited by a short timeframe and few databases. A systematic review process was not undertaken; rather it was a rapid review.124
Third, there are multiple publications for each of the trials included in the meta-analyses. Data extraction was difficult as time periods varied within and across reports of these studies; also the follow-up denominator was difficult to determine for several end points. Sample size denominators were not always consistent between papers written on the same trial. We also cannot be certain that there is no publication bias.

**Future Research**

In January 2000, the FDA approved the use of digital mammography in the United States. In September 2005, preliminary results from a large clinical trial that compared digital mammography to film mammography were published. The overall diagnostic accuracy of digital and film mammography as a means of screening for breast cancer is similar, but digital mammography is more sensitive despite similar specificity in women under the age of 50, women with radiographically dense breasts, and premenopausal or perimenopausal women. The clinical impact of this increase in sensitivity is unknown, but its magnitude suggests that it might be clinically relevant. Since the trials included in our systematic review used film mammography, determining whether digital mammography improves the benefit associated with screening (especially in younger women) would require further study.

Technologies such as breast MRI have not been adequately studied in the screening of average-risk women. Until the availability of MRI is improved and the cost decreased, it is unlikely to be a consideration for population screening. To date, there is published literature only for screening of high-risk individuals such as BRCA1 and BRCA2 mutation carriers. Current screening recommendations for breast MRI from the American Cancer Society are for documented BRCA carriers, first-degree relatives, or women with an estimated lifetime risk of breast cancer >25%. The specificity of MRI is significantly lower than that of mammography in all studies to date, resulting in more recalls and biopsies. Call-back rates for additional imaging ranged from 8% to 17% in the MRI screening studies, and biopsy rates ranged from 3% to 15%. Moreover, even though the sensitivity of MRI for detecting invasive breast cancers is high, 99%, its detection of noninvasive in situ cancers is lower (70%–90%), and mammography is still essential in order to detect microcalcifications that are not seen on MRI. Although several trials reported looking at the accuracy and positive predictive value of MRI and mammography in women with high breast density, all these trials have been conducted in women known or strongly suspected to have malignancies within the breast. To this point, no Phase III randomized trial reported has shown a reduction in either mortality or the size of diagnosed breast cancer when comparing breast MRI with mammography in women selected for high mammographic density alone.

Equity of access is an issue in Canada for Aboriginal women and women living in rural and remote areas. Removing barriers to access might benefit from exploratory qualitative research.

Most women value joint decision making with their primary care provider about breast cancer screening. Some guidelines propose that those discussions happen. However, healthcare providers may be uncomfortable with such discussions. Research is required to determine how best to engage in that discussion and how practitioners can provide a balanced perspective on the potential benefits and harms individualized for each woman.
Conclusion

This review found no new trials of the effectiveness of breast screening. Meta-analyses of mammography screening trials indicate that mammography significantly reduces breast cancer mortality among women aged 39–49 and 50–69. Although pooled results were nominally nonsignificant among women aged ≥70 years, there is insufficient evidence to conclude that screening is less effective in this subgroup. However, the absolute benefit of screening on breast cancer mortality was small in women of all ages and may be partially offset by harms related to false-positives and overdiagnosis. New technologies are advancing rapidly in the field of breast imaging, and future trials will be essential in assessing risk and benefit in screening the Canadian population.
Reference List


118. Northern Secretariat of the BC Centre of Excellence for Women's Health. The determinants of women's health in northern rural and remote regions: examples and recommendations from northern British Columbia. Prince George, BC: University of Northern British Columbia; 2010. Available at: http://unbc.ca/assets/northernfire/WmNorth.PDF.


List of Figures

Figure 1: Analytic Framework and Key Questions
Figure 2: Search Results
Figure 3: Preference and Values Search Results
Figure 1: Analytic Framework and Key Questions

Screening
a. Mammography (film and digital) or MRI for women aged 40–49 and ≥70
b. CBE alone and with mammography for women aged 40–49 and ≥70
c. BSE (all ages)

Average-risk
Women aged ≥40 without breast cancer

Reduction of late-stage invasive breast cancer

Harms of screening include:
- radiation exposure
- pain
- psychological responses
- false-positive and false-negative test results
- overdiagnosis

Reduced breast cancer mortality and all cause mortality

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

Contextual Questions:
1. What is the cost-effectiveness of screening?
2. What are patient preferences and values with regard to breast cancer screening?
3. What is the effectiveness of screening for specific subpopulations (rural and remote, Aboriginal, or other ethnic populations)?
4. What is the evidence of optimal frequency of screening with mammography?

Source: Nelson, 2009³
Figure 2: Search Results

920 Citations

220 Potentially Relevant

17 Included

920 Potentially Relevant Citations

700 Excluded at Title and Abstract Screening

105 Not Average-risk Population

6 Not about Mammography, CBE or BSE

53 Outcomes Not Mortality, Harms or Cost

39 Study Design

11 Harms (9 Primary Studies, 2 Systematic Reviews)

5 Cost-Effectiveness

1 Systematic Review (Mortality)
Figure 3: Preference and Values Search Results

662 Citations

636 Excluded at Title and Abstract Screening

26 Relevant at Title and Abstract Screening

3 Systematic Reviews

23 Primary Studies
Appendices

Appendix 1: Search Terms for Mammography, Harms and Costs
Appendix 2: Search Terms for Patient Preferences and Values
Appendix 3: Search Terms for Subpopulations
Appendix 4: Search Terms for Breast Cancer Frequency
Appendix 5: Grey Literature Search
Appendix 6: Characteristics of Included Studies
Appendix 7: Evidence Sets
Appendix 8: List of Studies Excluded at Full Text Screening
Appendix 9: List of External Reviewers – Protocol
Appendix 10: List of External Reviewers – Evidence Synthesis
Appendix 1: Search Terms for Mammography, Harms, and Costs

Screening:
OVID-Medline
October, 2010
1. exp breast neoplasms/
2. exp neoplasms/di
3. exp breast/
4. 2 and 3
5. 1 or 4
6. exp mass screening/
7. (screen$ or (routine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
8. 6 or 7
9. 5 and 8
10. exp physical examination/
11. exp breast/
12. exp breast neoplasms/
13. 11 or 12
14. 10 and 13
15. exp mammography/
16. 9 and 14
17. 9 and 15
18. exp mortality/
19. mo.fs.
20. 18 or 19
21. 16 and 20
22. 17 and 20
23. 21 or 22
24. limit 23 to (english language and humans)
25. limit 24 to (meta analysis or practice guideline or randomized controlled trial)
26. (random$ or rct).mp.
27. 24 and 26
28. (meta-analy$ or metaanaly$ or (systematic$ adj10 review$)).mp.
29. 24 and 28
30. 25 or 27 or 29
31. 24 not 30
32. limit 31 to ed=20081101-20100302
33. limit 30 to ed=20081101-20100302

Cochrane Central
October, 2010
1. ((breast$ or mammary) adj3 (neoplas$ or tumor$ or cancer$ or carcinom$)).mp.
2. (screen$ or (routine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
3. ((clinical$ or physical$) adj3 (exam$ or detect$ or diagnos$)).mp.
4. 2 or 3
5. 1 and 4
6. limit 5 to yr="2008 -Current"
Cochrane Database of Systematic Reviews

October, 2010
1. ((breast$ or mammary) adj3 (neoplas$ or tumor$ or cancer$ or carcinom$)).mp.
2. (screen$ or (routine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
3. ((clinical$ or physical$) adj3 (exam$ or detect$ or diagnos$)).mp.
4. 2 or 3
5. 1 and 4
6. limit 5 to last 2 years
7. ((breast$ or mammary) adj3 (neoplas$ or tumor$ or cancer$ or carcinom$)).kw.
8. 1 not 7
9. 4 and 7
10. limit 9 to last 2 years

Digital Mammography:

MERSC_DigitalBreastScreening_medline

October, 2010
1. exp breast neoplasms/
2. exp neoplasms/di
3. exp breast/
4. 2 and 3
5. 1 or 4
6. exp mass screening/
7. (screen$ or (rountine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
8. 6 or 7
9. 5 and 8
10. exp physical examination/
11. exp breast/
12. exp breast neoplasms/
13. 11 or 12
14. 10 and 13
15. exp mammography/
16. 9 and 14
17. 9 and 15
18. 16 or 17
20. exp image processing, computer-assisted/
21. exp mammography/
22. 20 and 21
23. 19 or 22
24. 8 and 23
25. limit 24 to english language
26. limit 25 to ed=20081101-20100302

Cochrane Central

October, 2010
1. ((digital$ or computer$) adj7 mammogra$).mp.
2. limit 1 to yr="2008 -Current"
Cochrane Database of Systematic Reviews

October, 2010
1. ((digital$ or computer$) adj7 mammogra$).mp.
2. limit 1 to yr="2008 -Current"

MRI:

Medline

October, 2010
1. exp breast neoplasms/
2. exp neoplasms/di
3. exp breast/
4. 2 and 3
5. 1 or 4
6. exp mass screening/
7. (screen$ or (routine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
8. 6 or 7
9. 5 and 8
10. exp physical examination/
11. exp breast/
12. exp breast neoplasms/
13. 11 or 12
14. 10 and 13
15. exp mammography/
16. 9 and 14
17. 9 and 15
18. 16 or 17
19. exp magnetic resonance imaging/
20. 5 and 19
21. 8 and 20
22. limit 21 to ed=20081101-20100302

Cochrane Central

October, 2010
1. ((breast$ or mammary) adj3 (neoplas$ or tumor$ or cancer$ or carcinom$)).mp.
2. (mri or magnetic resonance imag$).mp.
3. 1 and 2
4. limit 3 to yr="2008 -Current"

Cochrane Database of Systematic Reviews

October, 2010
1. ((breast$ or mammary) adj3 (neoplas$ or tumor$ or cancer$ or carcinom$)).mp.
2. (mri or magnetic resonance imag$).mp.
3. 1 and 2
4. limit 3 to yr="2008 -Current"
DCIS:

Medline

October, 2010
1. exp carcinoma, intraductal, noninfiltrating/
2. exp breast neoplasms/
3. 1 and 2
4. overdiagnos$.mp.
5. over-diagnos$.mp.
6. (overtreat$ or over-treat$).mp.
7. exp Diagnostic errors/
8. exp mass screening/
9. exp mammography/
10. 8 or 9
11. 3 and 7 and 10
12. 4 or 5 or 6
13. 3 and 12
14. limit 13 to ed=20081101-20100302

Adverse Effects:

Medline

October, 2010
1. exp mammography/
2. exp physical examination/
3. exp mass screening/
4. 1 or 2 or 3
5. exp breast/
6. exp breast diseases/di, ep
7. 5 or 6
8. 4 and 7
9. exp mammography/ae, ct
10. exp physical examination/ae, ct
11. exp mass screening/ae, ct
12. 9 or 10 or 11
13. 7 and 12
14. exp diagnostic errors/
15. (overtest$ or overdiagnos$ or over-test$ or over-diagnos$).mp.
16. misdiagnos$.mp.
17. (false$ adj (positiv$ or negativ$)).mp.
18. (((incorrect$ or false$ or wrong$ or bias$ or mistake$ or error$ or erroneous$) adj3 (result$ or finding$ or test$ or diagnos$)).mp.
19. (((inappropriate$ or unneccess$ or unnee$) adj3 (treat$ or Surg$ or therap$ or regimen$)).mp.
20. (observ$ adj3 bias$).mp.
21. or/14-20
22. 8 and 21
23. exp "wounds and Injuries"/ci, et
24. exp stress, psychological/
25. exp prejudice/
26. exp stereotyping/
27. or/23-26  
28. 8 and 27  
29. 13 or 22 or 28  
30. limit 29 to english language  
31. limit 30 to (meta analysis or randomized controlled trial)  
32. exp evaluation studies/  
33. comparative study.pt.  
34. exp epidemiologic studies/  
35. 32 or 33 or 34  
36. 30 and 35  
37. 31 or 36  
38. limit 37 to ed=20081101-20100302  

**Cochrane Central**  
October, 2010  
1. exp mammography/  
2. mammogra$.mp.  
3. exp physical examination/  
4. ((physical$ or clinical$ or manual$) adj3 exam$).mp.  
5. exp mass screening/  
6. screen$.mp.  
7. or/1-6  
8. exp breast/  
9. exp breast diseases/di, ep  
10. (breast$ or mammar$).mp.  
11. or/8-10  
12. 7 and 11  
13. ((advers$ adj3 effect$) or harm$ or contraindicat$).mp.  
14. ae.fs.  
15. or/13-14  
16. 12 and 15  
17. exp mammography/ae, ct  
18. exp physical examination/ae, ct  
19. exp mass screening/ae, ct  
20. or/17-19  
21. 11 and 20  
22. exp diagnostic errors/  
23. (overtest$ or overdiagnos$ or over-test$ or over-diagnos$).mp.  
24. (false$ adj (result$ or positiv$ or negativ$)).mp.  
25. (observ$ adj3 bias$).mp.  
26. (diagnos$ adj3 (error$ or mistak$ or incorrect$)).mp.  
27. or/22-26  
28. 12 and 27  
29. exp "wounds and Injuries"/ci, et  
30. exp stress, psychological/  
31. exp prejudice/  
32. exp stereotyping/  
33. (anxiet$ or anxious$ or fear$ or discriminat$ or unfair$ or prejudic$ or stigma$ or stereotyp$).mp.  
34. or/29-33  
35. 12 and 34
36. 16 or 21 or 28 or 35
37. limit 36 to yr="2008 -Current"

Cost:

Medline
October, 2010
1. exp breast neoplasms/
2. exp neoplasms/di
3. exp breast/
4. 2 and 3
5. 1 or 4
6. exp mass screening/
7. (screen$ or (rountine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
8. 6 or 7
9. 5 and 8
10. exp physical examination/
11. exp breast/
12. exp breast neoplasms/
13. 11 or 12
14. 10 and 13
15. exp mammography/
16. 9 and 14
17. 9 and 15
18. 16 or 17
19. exp "Costs and Cost Analysis"/
20. 18 and 19
21. limit 20 to english language
22. limit 21 to ed=20081101-20100302

Cochrane Central
October, 2010
1. ((breast$ or mammary) adj3 (neoplas$ or tumor$ or cancer$ or carcinom$)).mp.
2. (screen$ or (rountine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
3. ((clinical$ or physical$) adj3 (exam$ or detect$ or diagnos$)).mp.
4. (cost or costs or costing or economic$ or financial$).mp.
5. 1 and (2 or 3) and 4
6. limit 5 to yr="2008 -Current"

Cochrane Database of Systematic Reviews
October, 2010
1. ((breast$ or mammary) adj3 (neoplas$ or tumor$ or cancer$ or carcinom$)).mp.
2. (screen$ or (rountine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
3. ((clinical$ or physical$) adj3 (exam$ or detect$ or diagnos$)).mp.
4. (cost or costs or costing or economic$ or financial$).mp.
5. 1 and (2 or 3) and 4
6. limit 5 to yr="2008 -Current"
Appendix 2: Search Terms for Patient Preferences and Values

**EBSCO CINAHL**
S1 TI breast cancer screening
S2 (MH "Breast Neoplasms/DI")
S3 (MM "Mammography")
S4 S1 or S2 or S3
S5 (MM "Cancer Screening")
S6 (MM "Breast Neoplasms+")
S7 S5 and S6
S8 S4 or S7
S9 MM "Patient Compliance" or MM "Consumer Participation" or MH "Patient Satisfaction" or MH "Treatment Refusal" or MH "Consumer Satisfaction"
S10 TX women? N3 preference? or TX women? N3 acceptance or TX women? N3 satisfaction or TX women? N3 experience?
S11 TX consumer? N3 preference? or TX consumer? N3 acceptance or TX consumer? N3 satisfaction or TX consumer? N3 experience?
S12 TX consumer? N3 choice? or TX patient? N3 choice? or TX women* N3 choice?
S13 S9 or S10 or S11 or S12
S14 S8 and S13
S15 S8 and S13 [Limiters - Publication Year from: 2000-2010; Language: English, French]

**Ovid MEDLINE(R)**

May 7, 2010

---------------------------------------------
1 breast cancer screening.ti.
2 exp *Breast Neoplasms/di
3 exp *Mammography/
4 or/1-3
5 *mass screening/
6 exp *Breast neoplasms/
7 5 and 6
8 4 or 7
9 *"patient acceptance of healthcare"/ or *patient compliance/ or *patient participation/ or patient satisfaction/ or patient preference/ or *treatment refusal/
10 (women? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
11 (consumer? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
12 (patient? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
13 willingness to pay.tw.
14 ((conjoint or contingent) adj3 (valuation or analysis)).tw.
15 or/9-14
16 8 and 15
17 limit 16 to (english or french)
18 limit 17 to yr="2000 -Current"
Appendix 3: Search Terms for Subpopulations

Ovid MEDLINE(R)
<1950 to June Week 5 2010>

---------------------------------------------
1 exp breast neoplasms/
2 exp neoplasms/di
3 exp breast/
4 2 and 3
5 1 or 4
6 exp mass screening/
7 (screen$ or (rountine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
8 6 or 7
9 5 and 8
10 exp physical examination/
11 exp breast/
12 exp breast neoplasms/
13 11 or 12
14 10 and 13
15 exp mammography/
16 9 and 14
17 9 and 15
18 16 or 17
19 exp Breast Neoplasms/di, ep, eh, mo [Diagnosis, Epidemiology, Ethnology, Mortality]
20 18 or 19
21 Ethnic Groups/
22 ethnic*.ti.
23 Rural Health/
24 Rural Population/
25 rural health services/
26 (rural or remote).ti.
27 (geographic and disparity).ti.
28 Indians, North American/
29 first nations.tw.
30 native canadian?.tw.
31 (aboriginal? and canada).tw.
32 Jews/
33 Ashkenazi jew?.tw.
34 or/21-33
35 20 and 34
36 limit 35 to english language
37 limit 36 to (meta analysis or "review")
38 (systematic* adj review*).tw.
39 37 or 38
40 36 and 39
41 limit 40 to yr="2006 -Current"
Appendix 4: Search Terms for Breast Cancer Frequency

Medline

Aug 27, 2010

1. exp breast neoplasms/
2. exp neoplasms/di
3. exp breast/
4. 2 and 3
5. 1 or 4
6. exp mass screening/
7. (screen$ or (rountine$ adj3 (test$ or check$ or diagnos$ or detect$))).mp.
8. 6 or 7
9. 5 and 8
10. exp physical examination/
11. exp breast/
12. exp breast neoplasms/
13. 11 or 12
14. 10 and 13
15. exp mammography/
16. 9 and 14
17. 9 and 15
18. exp mortality/
19. mo.fs.
20. 18 or 19
21. 16 and 20
22. 17 and 20
23. 21 or 22
24. limit 23 to (english or french)
25. limit 24 to humans
26. (biannual or bi-annual).tw.
27. schedule.tw.
28. frequency.tw.
29. (interval not confidence interval).tw.
30. (annual* or yearly).tw.
31. biennial.tw.
32. 26 or 27 or 28 or 29 or 30 or 31
33. 25 and 32
34. limit 33 to yr="2000 -Current"
Appendix 5: Grey Literature Search

Google search limited to Canada

- “breast cancer screening AND harms”
- “mammography AND harms”
- “mammography AND costs”
- “breast cancer screening AND costs”

Specific Sites Search: search terms included “breast cancer screening” OR “mammography” OR “breast cancer”

The first set of sites was identified using CADTH’s Grey Matters: a practical search tool for evidence-based medicine.

- Agence d’évaluation des technologies et des modes d’intervention en santé (AETMIS), Québec
  http://www.aetmis.gouv.qc.ca/
- Canadian Agency for Drugs and Technologies in Health (CADTH)
  http://www.cadth.ca
- Centre for Evaluation of Medicines (Father Sean O’Sullivan Research Centre; St. Joseph’s Healthcare Hamilton; and McMaster University, Faculty of Health Sciences, Hamilton, Ontario)
  http://www.thecem.net/
- Centre for Health Services and Policy Research, University of British Columbia
  http://www.chspr.ubc.ca/cgi-bin/pub
- Health Quality Council, Saskatchewan
  http://www.hqc.sk.ca/
- Institute for Clinical Evaluative Sciences (ICES), Ontario
  http://www.ices.on.ca/
- IHE Institute of Health Economics, HTA Unit, Alberta
  http://www.ihe.ca/publications/library/
- Manitoba Centre for Health Policy (MCHP)
  http://umanitoba.ca/medicine/units/mchp/
- Ontario Health Technology Advisory Committee (OHTAC): Analyses and Recommendations
  http://www.health.gov.on.ca/english/providers/program/ohtac/tech/recommend/rec_mn.html
- Technology Assessment Unit of the McGill University Health Centre
  http://www.mcgill.ca/tau/publications/

Individual sites were then searched using the same terms.

- Canadian Cancer Society
  http://www.cancer.ca/
• Canadian Institute for Health Information (CIHI)
• Canadian Partnership Against Cancer
  http://www.partnershipagainstcancer.ca/
• Cancer Care Ontario
  http://www.cancer.ca/ontario
  Searched: Results Found
• Centre for Health Economics and Policy Analysis (CHEPA), McMaster University
  http://www.chepa.org/
• Health Canada
  http://www.hc-sc.gc.ca/english/
• Institute of Health Economics (IHE)
  http://www.ihe.ca
• Public Health Agency of Canada
• Statistics Canada
  http://www.statcan.gc.ca/start-debut-eng.html
### Appendix 6: Characteristics of Included Studies

<table>
<thead>
<tr>
<th>First Author</th>
<th>Andersson I¹⁵,¹⁷,¹²⁷,¹²⁸</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Sweden</td>
</tr>
<tr>
<td>Name of Study</td>
<td>Malmö mammographic screening trial (MMST1)</td>
</tr>
<tr>
<td>Objective</td>
<td>To determine whether mortality from breast cancer could be reduced by repeated mammographic screening.</td>
</tr>
</tbody>
</table>
| Methods      | **Design:** birth year cohorts of city population separately randomized into study and control groups  
                **Selection:** all women born 1908 to 1932 from population registry of Malmö  
                **Blinding:** validation of endpoints completed by blinded reviewers |
| Participants | **Sample:** screening clinic outside of main hospital  
                Women over 45 years; 21,088 invited for screening and 21,195 in control group.  
                **Characteristics:** 100% female  
                Age Range: 45 to 79 years  
                **Withdrawals/Drop-outs:**  
                Study Group:  
                74% round 1  
                70% rounds 2–5  
                Control Group: not reported  
                All subjects included had an endpoint recorded (alive or dead). |
| Intervention | Type of Mammography Screening Equipment: “up to date film screen mammography, improved equipment being used as it became available”  
                Timing of Intervention: 18 to 24 months for 5 rounds  
                Length of Follow-up: mean 8.8 years |
<table>
<thead>
<tr>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endpoint is mortality from breast cancer as the underlying cause of death as determined by a blinded independent committee.</td>
</tr>
<tr>
<td>Cause of death taken from national mortality registry.</td>
</tr>
</tbody>
</table>

**1993 results**\(^{71}\)

**Breast Cancer Deaths:**

- Study Group: 87  
  Control Group: 108  
  RR: 0.81, 95% CI 0.62–1.07

**1996 results**\(^{71}\)

**Breast Cancer Deaths:**

**Ages 45–70:**

- Study Group: 161  
  Control Group: 198  
  RR: 0.82, 95% CI 0.67–1.00

**Ages 50–59:**

- Study Group: 88  
  Control Group: 90  
  RR: 0.98, 95% CI 0.75–1.29

**Ages 60–69:**

- Study Group: 46  
  Control Group: 72  
  RR: 0.64, 95% CI 0.45–0.92

| Comments |
Appendix 6: Characteristics of Included Studies (continued)

<table>
<thead>
<tr>
<th>First Author</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bjurstam N⁶³,¹²⁹</td>
<td>Sweden (Göteborg)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of Study</th>
<th>The Göteborg Trial</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Objective</th>
<th>Determining whether mammographic screening results in a reduction in breast cancer mortality</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Methods</th>
<th>Design: Quasi-randomized</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• 1923–1935 birth year cohorts randomized by birth date (18%)</td>
</tr>
<tr>
<td></td>
<td>• 1936–1944 birth-year cohorts randomized individually (82%)</td>
</tr>
<tr>
<td>Selection:</td>
<td>Inclusion: women born between 1923 and 1944 who live in Göteborg, Sweden</td>
</tr>
<tr>
<td></td>
<td>Exclusion: women with a previous history of breast cancer were not included in the analysis</td>
</tr>
<tr>
<td></td>
<td>Blinding: underlying cause of death classified by an endpoint committee blinded from the study randomization.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participants</th>
<th>Sample:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Including Women Previously Diagnosed with Breast Cancer:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intervention Group: 21,904</td>
</tr>
<tr>
<td></td>
<td>Control Group: 30,318</td>
</tr>
<tr>
<td>Sample Used for Analysis:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intervention Group: 21,650</td>
</tr>
<tr>
<td></td>
<td>Control Group: 29,961</td>
</tr>
<tr>
<td>Ages 39–49:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intervention Group: 11,724</td>
</tr>
<tr>
<td></td>
<td>Control Group: 14,217</td>
</tr>
<tr>
<td>Ages 50–59:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intervention Group: 9,926</td>
</tr>
<tr>
<td></td>
<td>Control Group: 15,744</td>
</tr>
<tr>
<td>Attendance at screens – intervention group mean 78.7%</td>
<td></td>
</tr>
<tr>
<td>Attendance at sole screen – control group 72.1%</td>
<td></td>
</tr>
<tr>
<td>Characteristics:</td>
<td>100% female</td>
</tr>
<tr>
<td>Age Range:</td>
<td>39–59 years</td>
</tr>
<tr>
<td>Withdrawals/Drop-outs:</td>
<td>not clearly stated</td>
</tr>
<tr>
<td>Study Recruitment Years:</td>
<td>1982 to 1984</td>
</tr>
</tbody>
</table>
| Intervention | Two view mammography used at the first round, single view mammography used after that unless single view was inappropriate due to density of the breast.  
**Type of Mammography Screening Equipment:** Kodak Min R imaging system used in first round; CGR Senograph 500 T unit used in subsequent rounds  
**Timing of Intervention:** 18 month intervals between screens.  
**Length of Follow-up:** up to 14 years |
| Outcomes | Three models were used to determine mortality  
**Endpoint Committee Model:** subjects with breast carcinoma diagnosed during the screening phase, two independent endpoint committees classified the underlying cause of death, identified through the Swedish Cancer Register, backed up by the Swedish Cause of Death register  
**Mortality:**  
**Ages 39–49:**  
Intervention Group: 25/11,724  
Control Group: 46/14,217  
RR: 0.65, 95% CI 0.40–1.05  
**Ages 50–59:**  
Intervention Group: 38/9,926  
Control Group: 66/15,744  
RR: 0.91, 95% CI 0.61–1.36  
**Swedish Cause of Death Evaluation Model:** subjects with breast carcinoma diagnosed during the screening phase, identified through the Swedish Cancer Register, backed up by the Swedish Cause of Death register  
**Mortality:**  
**Ages 39–49:**  
Intervention Group: 23/11,724  
Control Group: 49/14,217  
RR: 0.56, 95% CI 0.34–0.91  
**Ages 50–59:**  
Intervention Group: 40/9,926  
Control Group: 68/15,744  
RR: 0.93, 95% CI 0.63–1.38  
**Swedish Cause of Death Follow-up Model:** outcomes as determined by data from the Swedish National Cause of Death Register up to December 31, 1996 (NB, these data were analyzed in this review) |
<table>
<thead>
<tr>
<th>Mortality:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ages 39–49:</td>
</tr>
<tr>
<td>Intervention Group: 34/11,724</td>
</tr>
<tr>
<td>Control Group: 59/14,217</td>
</tr>
<tr>
<td>RR: 0.69, 95% CI 0.45–1.05</td>
</tr>
<tr>
<td>Ages 50–59:</td>
</tr>
<tr>
<td>Intervention Group: 54/9,926</td>
</tr>
<tr>
<td>Control Group: 103/15,744</td>
</tr>
<tr>
<td>RR: 0.83, 95% CI 0.60–1.15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women with a previous history of breast cancer were included in the study but were not included in the analysis.</td>
</tr>
<tr>
<td>The study group was invited to a screening every 18 months. The control group received an invitation to a one-time screening at the end of the study.</td>
</tr>
<tr>
<td>Trial closed one round earlier in women older than 50 years due to the introduction of a local policy of routine service screening for women in this age group.</td>
</tr>
</tbody>
</table>
### Appendix 6: Characteristics of Included Studies (continued)

<table>
<thead>
<tr>
<th>First Author</th>
<th>Frisell J&lt;sup&gt;64,71&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Sweden</td>
</tr>
<tr>
<td><strong>Name of Study</strong></td>
<td>Stockholm Mammographic Screening Trial</td>
</tr>
<tr>
<td><strong>Objective</strong></td>
<td>To determine whether mammographic screening would lead to a reduction in mortality from breast cancer.</td>
</tr>
</tbody>
</table>
| **Methods** | **Design:** Quasi-randomized, selected by birth dates, Study Population (SP) born 1<sup>st</sup> through 10<sup>th</sup> and 21<sup>st</sup> through 31<sup>st</sup>; Control Population (CP) born 11<sup>th</sup> through 20<sup>th</sup> day of the month  
**Selection:** residents of southeast Stockholm, Sweden  
**Blinding:** unclear |
| **Participants** | **Sample:** 60,261  
40,318 study (SP); 19,943 control (CP)  
**Characteristics:** 100% female  
Age Range: 40 to 64 years at recruitment  
**Withdrawals/Drop-outs:** attendance rate was 81% and 80% in the respective rounds for SP; 77% in the screening of the CP  
**Study Recruitment Years:** 1981 to 1985 |
| **Intervention** | Single view mammography vs. usual care  
**Type of Mammography Screening Equipment:** CGR Mammography (Senograph 500T)  
**Timing of Intervention:** approximately 2 years between mammograms  
**Length of Follow-up:** Mean 11.4 years |
| **Outcomes** | Endpoint in the trial was breast cancer death – defined as “death with breast cancer present at death (locoregional or distant disease)” Causes of death were assessed by an independent committee after a review of all total breast cancer cases.  
**Breast Cancer Deaths 50 to 64 Years:**  
SP: 48/24,836  
CP: 37/12,957  
RR: 0.65, 95% CI 0.45–94 |
<table>
<thead>
<tr>
<th></th>
<th>Breast Cancer Deaths 40 to 49 Years:</th>
<th>All Breast Cancer Deaths:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SP: 34/14,303</td>
<td>SP: 82/39,139</td>
</tr>
<tr>
<td></td>
<td>CP: 13/8,021</td>
<td>CP: 50/20,978</td>
</tr>
<tr>
<td></td>
<td>RR: 1.47, 95% CI 0.77–2.78</td>
<td>RR: 0.88, 95% CI 0.62–1.25</td>
</tr>
</tbody>
</table>

**Comments**

Long screening intervals, the use of single-view mammography and the fact that more than 50% of the women in age group 40–59 years were still below 50 years of age when the study was closed, were all factors that could have influenced the results in the age group 40 to 49 years.

The reporting of the numbers varies between publications.
<table>
<thead>
<tr>
<th>First Author</th>
<th>Habbema JDF[^66,130-133]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>USA</td>
</tr>
<tr>
<td>Name of Study</td>
<td>Health Insurance Plan (HIP), New York</td>
</tr>
<tr>
<td>Objective</td>
<td>The first breast cancer screening trial, which was initiated in December 1963 to explore the efficacy of mammography screening. Breast cancer and mortality from breast cancer were examined by treatment group (study vs. control) and by entry-age subgroup</td>
</tr>
</tbody>
</table>
| Methods      | **Design:** Random assignment to study and control groups – matched by age, gender and size of family group  
**Selection:** New York Health Plan members: 31,092 were invited and 64% of sample agreed to come for screening (n=20,211)[^130]  
**Blinding:** blinded review of death certificate and medical records |
| Participants | **Sample:** enrollees in the Health Insurance Plan (HIP) of Greater New York  
Study Group: 30,245  
Control Group: 30,245  
**Characteristics:** 100% female  
Age Range: 40 to 64 years  
**Withdrawals/Drop-outs:** not reported  
**Study Recruitment Years:** 1963 from the registry of the Health Insurance Plan of New York with 23 of its 31 affiliated medical groups |
| Intervention | Study group women were invited for screening, an initial examination, and three annual re-examinations. Screening consisted of film mammography (cephalocaudal and lateral views of each breast) and clinical examination of breasts.  
**Timing of Intervention:** screening at annual intervals for 3 years  
**Length of Follow-up:**  
Longest follow-up: 18 years  
Median: 16 years |
| Outcomes     | Death with breast cancer as the underlying cause according to internationally accepted rules. Only deaths occurring among breast cancer cases diagnosed within 7 years after entry in the study are taken into account. The study group had about 25% lower breast cancer mortality among women aged 40–49 and 50–59 at time of entry than did the control group. |
### Breast Cancer Mortality Ages 40 to 64:
- **Study Group:** 165/30,245
- **Control Group:** 212/30,245

### Breast Cancer Mortality Ages 40 to 49:
- **Study Group:** 64/13,740*
- **Control Group:** 82/13,740*
- **RR:** 0.78, 95% CI 0.56–1.08

### Breast Cancer Mortality Ages 50 to 59:
- **Study Group:** 77/12,855*
- **Control Group:** 97/12,855*
- **RR:** 0.79, 95% CI 0.59–1.07

### Breast Cancer Mortality Ages 60 to 64:
- **Study Group:** 24/3,650*
- **Control Group:** 33/3,650*
- **RR:** not reported

| Comments | To a large extent the difference among the 40- to 49-year-olds occurred in the subgroup with breast cancer diagnosed after these women had passed their 50th birthday, and utility of screening women in their forties is questionable. The reporting of the sample size of the actual study population varies between publications. |

* Denominators are estimated. Study doesn’t clearly state number of participants but rather states that “the numbers are about the same in study and control groups.” The numbers used in this table are the same as those used by the USPSTF 2002 report.
Appendix 6: Characteristics of Included Studies (continued)

<table>
<thead>
<tr>
<th>First Author</th>
<th>Miller AB^56,59,134</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Canada</td>
</tr>
<tr>
<td>Name of Study</td>
<td>CNBSS-1</td>
</tr>
</tbody>
</table>

**Objective**
To evaluate the efficacy of the combination of annual screening with mammography, physical examination of the breasts (CBE) and the teaching of breast self-examination (BSE) in reducing the rate of death from breast cancer among women aged 40 to 49 years on entry.

**Methods**
- **Design:** individually randomized controlled trial – mammography + physical exam (MP) vs. annual physical exam only (PO)
- **Selection:** female volunteers with no history of breast cancer and no mammography in the previous 12 months
- **Blinding:** the examiners did not have access to the group assignments. Analysis of data by several reviewers

**Participants**
- **Sample:** 50,430 women enrolled
  - Intervention Group: 25,214
  - Control Group: 25,216
- **Characteristics:** 100% female
  - Mean Age: not reported
  - Age Range: 40 to 49
- **Withdrawals/Drop-outs:** 42 distributed equally between two groups were excluded from the analysis
- **Study Recruitment Years:** 1980 to 1985

**Intervention**
Annual two view mammography and annual physical examination for 4 to 5 years
- **Timing of Intervention:** annually for 5 years
- **Length of Follow-up:** 11 to 16 years; mean 8.5 years

**Outcomes**
- Death due or probably due to breast cancer.
- All diagnoses of breast cancer were histologically verified.
- All cause of death only reported to 1993.

**Breast Cancer Mortality (Ages 40–49):**
- **To June 30, 1996:**
  - MP: 105/25,214
  - PO: 108/25,216
  - RR: 97, 95% CI 0.74-1.27

**Comments**
### Appendix 6: Characteristics of Included Studies (continued)

<table>
<thead>
<tr>
<th>First Author</th>
<th>Miller AB&lt;sup&gt;57,60,134&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Canada</td>
</tr>
<tr>
<td>Name of Study</td>
<td>CNBSS-2</td>
</tr>
<tr>
<td><strong>Objective</strong></td>
<td>To evaluate the efficacy of the combination of annual screening with mammography, physical examination of the breasts (CBE) and the teaching of breast self-examination (BSE) in reducing the rate of death from breast cancer among women aged 50–59 years on entry.</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td><strong>Design:</strong> individually randomized controlled trial – mammography + physical exam (MP) vs. annual physical exam only (PO)  \  <strong>Selection:</strong> female volunteers with no history of breast cancer and no mammography in the previous 12 months  \  <strong>Blinding:</strong> the examiners did not have access to the group assignments. Analysis of data by several reviewers</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
<td><strong>Sample:</strong> 39,405 women  \  MP: 19,711  \  PO: 19,694  \  <strong>Characteristics:</strong> Female  \  Mean Age: not reported  \  Age Range: 50 to 59  \  <strong>Withdrawals/Drop-outs:</strong> 71 distributed equally between two groups were excluded from the analysis  \  <strong>Study Recruitment Years:</strong> 1980–1985</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td>Annual two view mammography and annual physical examination for 4 to 5 years.  \  <strong>Timing of Intervention:</strong> annually for 5 years  \  <strong>Length of Follow-up:</strong> mean 13 years (range 11.3 to 16 yrs)</td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td><strong>Breast Cancer Mortality:</strong>  \  To June 30, 1996:  \  MP: 107/19,711  \  PO: 105/19,694  \  RR. 1.02, 95% CI 0.78-1.33</td>
</tr>
<tr>
<td><strong>Comments</strong></td>
<td></td>
</tr>
</tbody>
</table>

---

67
# Appendix 6: Characteristics of Included Studies (continued)

<table>
<thead>
<tr>
<th>First Author Country</th>
<th>Moss S\textsuperscript{61,135} UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of Study</td>
<td>AGE</td>
</tr>
<tr>
<td>Objective</td>
<td>To study the effect on mortality of inviting women for annual mammography from age 40</td>
</tr>
</tbody>
</table>
| Methods              | **Design:** individual randomization by computer; randomization ratio 1:2  
                        **Selection:** women aged 39 to 41 from the health authority general practices  
                        **Blinding:** cause of death was taken from death certificate |
| Participants         | **Sample:** 160,840 randomized  
                        Intervention Group: 53,884  
                        Control Group: 106,956  
                        **Characteristics:** 100% female  
                        Age Range: 39 to 41 at recruitment  
                        **Withdrawals/Drop-outs:** none reported  
                        **Study Recruitment Dates:** 1991 to 1997 |
| Intervention         | Annual mammography screening to age 48 years vs usual care  
                        **Type of Mammography:** two view mammography at first screen, single view mammography after the initial screen with recall for full assessment if an abnormality suspected  
                        **Timing of Intervention:** annually  
                        **Length of Follow-up:** Mean 10.7 years; Range 7 to 14 years |
| Outcomes             | Intention to Treat (ITT) Analysis  
                        **Breast Cancer Mortality All Ages:**  
                        Intervention Group: 105/53,884  
                        Control Group: 251/106,956  
                        RR: 0.83, 95% CI 0.66–1.04 |
| Comments             |                                   |
### Appendix 6: Characteristics of Included Studies (continued)

<table>
<thead>
<tr>
<th>First Author</th>
<th>Tabár L\textsuperscript{62,65,136}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Sweden, Kopparberg (W-County)</td>
</tr>
</tbody>
</table>

| Name of Study      | Swedish Two-County Trial (W-County) |

| Objective          | Comparison of mortality between an invitation to screening group, and a control group not invited |

<table>
<thead>
<tr>
<th>Methods</th>
<th>Design: Cluster randomized by geographical area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection:</td>
<td>Inclusion: Women between 40 and 74 in Kopparberg (W-County)</td>
</tr>
<tr>
<td></td>
<td>Exclusion: Women without a permanent address, women diagnosed with breast cancer before randomization.</td>
</tr>
<tr>
<td></td>
<td>Blinding: unclear</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participants</th>
<th>Sample:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Active Study Population (ASP): 38,598</td>
</tr>
<tr>
<td></td>
<td>Passive Study Population (PSP): 18,582</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristics:</th>
<th>Age Range: 40 to 74 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Withdrawals/Drop-outs:</td>
<td>The 70- to 74-year-old cohort was discontinued after 2\textsuperscript{nd} screening due to low response rate.</td>
</tr>
</tbody>
</table>

| Study Recruitment Years: | 1978 to 1985 |

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Single view screen film mammography</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Mammography Screening Equipment:</td>
<td>not reported.</td>
</tr>
</tbody>
</table>

| Timing of Intervention: | Ages 40–49 on average every 24 months |
|                        | Ages 50–74 on average every 33 months       |

| Length of Follow-up: | 13 years stated in the methods – up to 20 years |

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Primary Outcome: Mortality from breast cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcomes from Tabár, 1995:</td>
<td></td>
</tr>
<tr>
<td>Ages 40 to 49:</td>
<td>ASP: 22/9,582</td>
</tr>
<tr>
<td></td>
<td>PSP: 16/5,031</td>
</tr>
</tbody>
</table>
| Ages 50 to 59: | ASP: 34/11,728  
PSP: 34/5,557 |
| Ages 60 to 69: | ASP: 44/11,973  
PSP: 35/5,555 |
| Ages 70 to 74: | ASP: 26/5,306  
PSP: 19/2,439 |

**Comments**

Women from towns/parishes/municipalities indicating the possibility of a homogenous population

Women from certain towns may all be systematically exposed to a certain factor (carcinogen etc.) that could affect results.

Numbers vary between publications.
## Appendix 6: Characteristics of Included Studies (continued)

<table>
<thead>
<tr>
<th>First Author</th>
<th>Tabár L62,65,71,136</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Sweden Östergötland (E-County)</td>
</tr>
</tbody>
</table>

| Name of Study | Swedish Two-County Trial (E-County) |

| Objective | Comparison of mortality between an invitation to screening group, and a control group not invited |

| Methods | **Design:** Cluster randomized by geographical area  
**Selection:** inclusion/exclusion  
Inclusion: Women in Östergötland, Sweden (E-County).  
Exclusion: Women without a permanent address, women diagnosed with breast cancer before randomization.  
**Blinding:** unclear |

| Participants | **Sample:**  
Tabár, 199562  
Active Study Population (ASP): 38,491  
Passive Study Population (PSP): 37,403  
Nyström, 200271:  
Intervention Group: 38,942  
Control Group: 37,675  
**Characteristics:**  
Age Range: 40 to 74 years  
Withdrawals/Drop-outs: 70- to 74-year-old cohort was discontinued after 2nd screening due to low response rate.  
Study Recruitment Years: 1978 to 1985 |

| Intervention | Single view screen film mammography  
**Type of Mammography Screening Equipment:** not reported.  
**Timing of Intervention:**  
Ages 40–49: on average every 24 months  
Ages 50–74: on average every 33 months  
**Length of Follow-up:** 13 to 20 years |
| Outcomes | Primary Outcome: Mortality
Breast Cancer Mortality: Nyström, 2002\textsuperscript{1}:

All Ages:  
  ASP: 177/38,942  
  PSP: 190/37,675

Ages 39 to 49:  
  ASP: 31/10,285  
  PSP: 30/10,459

Ages 50 to 59:  
  ASP: 53/12,011  
  PSP: 54/11,495

Ages 60 to 69:  
  ASP: 64/11,573  
  PSP: 83/10,862 |
| Comments | Women from towns/parishes/municipalities indicating the possibility of a homogenous population
Women from certain towns may all be systematically exposed to a certain factor (carcinogen etc.) that could affect results.
Numbers vary between publications. |
Appendix 7: Evidence Sets

Evidence Set 1: KQ1a – Breast Cancer Mortality (Aged 40–49)
Evidence Set 2: KQ1a – Breast Cancer Mortality (Aged 70–74)
Evidence Set 3: KQ1a – Breast Cancer Mortality (Aged 50–69)
Evidence Set 4: KQ1a – Breast Cancer Mortality (Aged 50–59)
Evidence Set 5: KQ1a – Breast Cancer Mortality (Aged 60–69)
Evidence Set 6: KQ1a – Breast Cancer Mortality (All Age Groups)
Evidence Set 7: KQ1a – All Cause Mortality
Evidence Set 8: KQ1c – BSE and All Cause Mortality (Aged ≥40)
Evidence Set 9: KQ2a – Harms of Screening/Mammography (Unnecessary Surgeries)
Evidence Set 10: CQ4 – Optimal Mammography Screening Intervals
## Risk of Bias for KQ1a (All Age Groups)

<table>
<thead>
<tr>
<th>Author/Year</th>
<th>Randomization</th>
<th>Allocation Concealment</th>
<th>Blinding</th>
<th>Loss to Follow-up /Intention to Treat Analysis (ITT) Principle Observed or per Protocol Analysis</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersson 1988</td>
<td>Individual randomization used in Malmö 1.</td>
<td>Not described</td>
<td>Validation of endpoints completed by blinded reviewers.</td>
<td>Higher attendance rate in first round (74%). - subsequent rounds (70%) - higher in younger than older women</td>
<td></td>
</tr>
<tr>
<td>Nyström 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malmö 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Göteborg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frisell 1997</td>
<td>Quasi-randomized Cluster sample – selection done individually based on date of birth.</td>
<td>Not described</td>
<td>unclear</td>
<td>Loss to follow-up rate is not clear. - all breast cancer deaths checked against the population register to ensure completeness of follow-up</td>
<td></td>
</tr>
<tr>
<td>Nyström 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stockholm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habbema 1986</td>
<td>Two random samples of women were selected and age and family-size stratified – process not described.</td>
<td>Not described</td>
<td>Blinded review of death certificate and medical records.</td>
<td>The assumption was made based on sample size in each group - actual numbers not provided for denominator - no relative risks or p values provided</td>
<td>Study conducted more than 30 years ago – USPSTF did not include an analysis on the basis of incompatible equipment.</td>
</tr>
<tr>
<td>HIP New York</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author/Year</td>
<td>Randomization</td>
<td>Allocation Concealment</td>
<td>Blinding</td>
<td>Loss to Follow-up /Intention to Treat Analysis (ITT) Principle Observed or per Protocol Analysis</td>
<td>Other</td>
</tr>
<tr>
<td>------------</td>
<td>---------------</td>
<td>-------------------------</td>
<td>---------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Miller 1992</td>
<td>Randomization was individual and stratified by centre and 5-year age group.</td>
<td>YES – blocking of the lists was unknown to the staff at the screening centre. Computer generated random numbers done by blocks.</td>
<td>The examiner had no access to the allocation list. Verification of cause of death blinded to group assignment.</td>
<td>ITT analysis - 1992 paper describes those women excluded from analysis</td>
<td>Mortality based on death certificates – linkage to the Canadian Mortality Database.</td>
</tr>
<tr>
<td>Moss 61,135</td>
<td>Ages 39–41 were randomly assigned in the ratio 1:2 to an intervention and control group.</td>
<td>YES - allocation to trial group was carried out on the health authority computer system using randomization software</td>
<td>Cause of death was taken from death certificate.</td>
<td>Less than 1% loss to follow-up. ITT analysis.</td>
<td>Control group members were unaware of their participation in the trial.</td>
</tr>
<tr>
<td>Tabár 2000</td>
<td>Cluster – by geographic area, traditional mechanical methods (flipping a coin).</td>
<td>Not described</td>
<td>Unclear</td>
<td>70 to 74 cohort discontinued after 2nd round of screening because of poor response rate but retained for intention to treat analysis of the trial mortality results.</td>
<td></td>
</tr>
<tr>
<td>Nyström 2002</td>
<td>Cluster – by geographic area, traditional mechanical methods (flipping a coin)</td>
<td>Not described</td>
<td>Unclear</td>
<td>70 to 74 cohort discontinued after 2nd round of screening because of poor response rate but retained for intention to treat analysis of the trial mortality results.</td>
<td></td>
</tr>
</tbody>
</table>
Evidence Set 1: KQ1a – Breast Cancer Mortality (Aged 40–49)

Does screening with mammography (film and digital) decrease breast cancer mortality for women aged 40–49?

- GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 40–49)
- Summary of Findings Table for KQ1a – Breast Cancer Mortality (Aged 40–49)
- Forest Plot for KQ1a – Breast Cancer Mortality (Aged 40–49)
### GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 40–49)

#### Studies included:
- **a:** HIP; Habbema et al.66
- **b:** Kopparberg (W-County); Tabár et al.62
c1: CNBSS-1; Miller et al.56
- **c:** Malmö; Nyström et al.71
e: Stockholm; Nyström et al.71
- **f:** Östergötland (E-County); Nyström et al.71
g: Göteborg; Bjurstam et al.63
- **h:** AGE; Moss et al.61

#### Quality Assessment

<table>
<thead>
<tr>
<th>No of Studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other Considerations</th>
<th>Screening</th>
<th>Control</th>
<th>Relative (95% CI)</th>
<th>Absolute per Million (Range)</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td><em><em>Breast Cancer Mortality for Ages 39-49</em>, All Trials (follow-up overall median 11.4 years)</em>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 (a-h)</td>
<td>randomized trials¹</td>
<td>serious²</td>
<td>no serious inconsistency³</td>
<td>no serious indirectness⁴</td>
<td>no serious imprecision⁵</td>
<td>none⁶</td>
<td>448/152,300 (0.2942%)</td>
<td>625/195,919 (0.3190%)</td>
<td>RR 0.8513 (0.7518 to 0.9639)</td>
<td>474 fewer (from 115 fewer to 792 fewer)</td>
<td>⊕⊕⊝ ⊦ MODERATE</td>
<td>CRITICAL</td>
</tr>
<tr>
<td><em><em>Breast Cancer Mortality for Ages 39-49</em>, 3 Truly Randomized Trials (follow-up overall median 11.4 years)</em>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (c1,d,h)</td>
<td>randomized trials</td>
<td>no serious risk of bias⁷</td>
<td>no serious inconsistency⁷</td>
<td>no serious indirectness⁴</td>
<td>no serious imprecision⁵</td>
<td>none⁶</td>
<td>263/92,666 (0.2838%)</td>
<td>425/144,451 (0.2942%)</td>
<td>RR 0.8545 (0.7308 to 0.9991)</td>
<td>428 fewer (from 3 fewer to 794 fewer)</td>
<td>⊕⊕⊕⊕ ⊦ HIGH</td>
<td>CRITICAL</td>
</tr>
<tr>
<td><em><em>Breast Cancer Mortality for Ages 39-49</em>, 5 Quasi-Randomized Studies (follow-up overall median 11.4 years)</em>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 (a,b,e,f,g)</td>
<td>randomized trials</td>
<td>serious⁸</td>
<td>no serious inconsistency⁹</td>
<td>no serious indirectness⁴</td>
<td>no serious imprecision⁵</td>
<td>none⁶</td>
<td>185/59,634 (0.3102%)</td>
<td>200/51,468 (0.3886%)</td>
<td>RR 0.8558 (0.6791 to 1.0784)</td>
<td>560 fewer (from 1,247 fewer to 305 more)</td>
<td>⊕⊕⊕ ⊦ MODERATE</td>
<td>CRITICAL</td>
</tr>
<tr>
<td><em><em>Breast Cancer Mortality for Ages 39-49</em>, Excludes HIP Study</em>* ⁸ (follow-up overall median 11.4 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 (b-h)</td>
<td>randomized trials</td>
<td>serious¹⁰</td>
<td>no serious inconsistency¹¹</td>
<td>no serious indirectness⁴</td>
<td>no serious imprecision⁵</td>
<td>none⁶</td>
<td>384/138,560 (0.2771%)</td>
<td>543/182,179 (0.2981%)</td>
<td>RR 0.8642 (0.7521 to 0.9930)</td>
<td>405 fewer (from 21 fewer to 739 fewer)</td>
<td>⊕⊕⊕ ⊦ MODERATE</td>
<td>CRITICAL</td>
</tr>
<tr>
<td><em><em>Breast Cancer Mortality for Ages 39-49</em>, 4 Quasi-Randomized Studies (Excludes HIP Study</em>*) (follow-up overall median 11.4 years)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (b,e,f,g)</td>
<td>randomized trials</td>
<td>serious⁸</td>
<td>no serious inconsistency¹²</td>
<td>no serious indirectness⁴</td>
<td>serious¹³</td>
<td>none⁶</td>
<td>121/45,894 (0.2637%)</td>
<td>118/37,728 (0.3128%)</td>
<td>RR 0.9089 (0.6554 to 1.2605)</td>
<td>285 fewer (from 1,078 fewer to 815 more)</td>
<td>⊕⊕⊝⊝ ⊦ LOW</td>
<td>CRITICAL</td>
</tr>
</tbody>
</table>

¹ five quasi-randomized and three truly randomized  
² blinding and concealment were not clear for five studies, so only three trials are considered truly randomized  
³ no heterogeneity exists; p-value for testing heterogeneity is 0.48 and I²=0%  
⁴ the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome  
⁵ total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
insufficient number of studies to assess publication bias
no heterogeneity exists; p-value for testing heterogeneity is 0.42 and $I^2=0$
concealment and blinding are not clear
no heterogeneity exists; p-value for testing heterogeneity is 0.31 and $I^2=17$
blinding and concealment were not clear for four studies, so only three trials are considered truly randomized
no heterogeneity exists; p-value for testing heterogeneity is 0.40 and $I^2=4$
no heterogeneity exists; p-value for testing heterogeneity is 0.22 and $I^2=33$
total sample size is large but the total number of events is <300 (a threshold rule-of-thumb value)
*The available data were based on women aged 39-49 although the focus of the review was for those aged 40-49.
**HIP was excluded because of the age of the study and the equipment used.
## Summary of Findings Table for KQ1a – Breast Cancer Mortality (Aged 40–49)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative Comparative Risks* (95% CI)</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed Risk, Number per Million Control</td>
<td>Corresponding Risk, Number per Million Screening</td>
<td>RR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 39–49*, All Trials</td>
<td>3,190</td>
<td>2,716</td>
<td>RR 0.8513 (0.7518 to 0.9639)</td>
<td>348,219 (8 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>(2,398 to 3,075)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 39–49*, 3 Truly Randomized Trials</td>
<td>2,942</td>
<td>2,514</td>
<td>RR 0.8545 (0.7308 to 0.9991)</td>
<td>237,117 (3 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>(2,150 to 2,940)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 39–49*, 5 Quasi-Randomized Studies</td>
<td>3,886</td>
<td>3,326</td>
<td>RR 0.8558 (0.6791 to 1.0784)</td>
<td>111,102 (5 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>(2,639 to 4,191)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 39–49*, Excludes HIP Study**</td>
<td>2,981</td>
<td>2,576</td>
<td>RR 0.8642 (0.7521 to 0.9930)</td>
<td>320,739 (7 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>(2,242 to 2,960)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 39–49*, 4 Quasi-Randomized Studies (Excludes HIP Study**)</td>
<td>3,128</td>
<td>2,843</td>
<td>RR 0.9089 (0.6554 to 1.2605)</td>
<td>83,622 (4 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>(2,050 to 3,942)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The assumed risk is the median control group risk. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence Interval; RR: Risk Ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1 five quasi-randomized and three truly randomized
2 blinding and concealment were not clear for five studies, so only three trials are considered truly randomized
3 no heterogeneity exists; p-value for testing heterogeneity is 0.48 and $I^2=0%$
4 the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
5 total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
6 insufficient number of studies to assess publication bias
7 no heterogeneity exists; p-value for testing heterogeneity is 0.42 and $I^2=0%$
8 concealment and blinding are not clear
9 no heterogeneity exists; p-value for testing heterogeneity is 0.31 and $I^2=17%$
10 blinding and concealment were not clear for four studies, so only three trials are considered truly randomized
11 no heterogeneity exists; p-value for testing heterogeneity is 0.40 and $I^2=4%$
12 no heterogeneity exists; p-value for testing heterogeneity is 0.22 and $I^2=3%$
13 total sample size is large but the total number of events is <300 (a threshold rule-of-thumb value)

*The available data were based on women aged 39–49 although the focus of the review was for those aged 40–49.

**HIP was excluded because of the age of the study and the equipment used.
## Forest Plot for KQ1a – Breast Cancer Mortality for Ages 40–49, All Trials

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental Events</th>
<th>Experimental Total</th>
<th>Control Events</th>
<th>Control Total</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bjurstam (Göteborg) 2003</td>
<td>34</td>
<td>11,724</td>
<td>59</td>
<td>14,217</td>
<td>8.7%</td>
<td>0.6988 [0.4586, 1.0649]</td>
<td></td>
</tr>
<tr>
<td>Habbema (HIP) 1986</td>
<td>64</td>
<td>13,740</td>
<td>82</td>
<td>13,740</td>
<td>14.5%</td>
<td>0.7805 [0.5633, 1.0814]</td>
<td></td>
</tr>
<tr>
<td>Miller (CNBSS-1) 2002</td>
<td>105</td>
<td>25,214</td>
<td>108</td>
<td>25,216</td>
<td>21.5%</td>
<td>0.9723 [0.7437, 1.2712]</td>
<td></td>
</tr>
<tr>
<td>Moss (AGE) 2006</td>
<td>105</td>
<td>53,884</td>
<td>251</td>
<td>106,956</td>
<td>29.8%</td>
<td>0.8303 [0.6614, 1.0425]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Malmö) 2002</td>
<td>53</td>
<td>13,568</td>
<td>66</td>
<td>12,279</td>
<td>11.9%</td>
<td>0.7267 [0.5067, 1.0424]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Östergötland) 2002</td>
<td>31</td>
<td>10,285</td>
<td>30</td>
<td>10,459</td>
<td>6.1%</td>
<td>1.0508 [0.6366, 1.7346]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Stockholm) 2002</td>
<td>34</td>
<td>14,303</td>
<td>13</td>
<td>8,021</td>
<td>3.8%</td>
<td>1.4667 [0.7745, 2.7775]</td>
<td></td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>22</td>
<td>9,582</td>
<td>16</td>
<td>5,031</td>
<td>3.7%</td>
<td>0.7219 [0.3795, 1.3734]</td>
<td></td>
</tr>
</tbody>
</table>

Total (95% CI) 152,300 195,919 100.0% 0.8513 [0.7518, 0.9639]

Total events 448 625

Heterogeneity: Tau² = 0.00; Chi² = 6.56, df = 7 (P = 0.48); I² = 0%

Test for overall effect: Z = 2.54 (P = 0.01)

(M-H: Mantel-Haenszel)
Evidence Set 2: KQ1a – Breast Cancer Mortality (Aged 70–74)

Does screening with mammography (film and digital) decrease breast cancer mortality for women aged 70–74?

- GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 70–74)
- Summary of Findings Table for KQ1a – Breast Cancer Mortality (Aged 70–74)
- Forest Plot for KQ1a – Breast Cancer Mortality (Aged 70–74)
GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 70–74)

Studies included:
b: Kopparberg (W-County); Tabár et al.62  f: Östergötland (E-County); Tabár et al.62

<table>
<thead>
<tr>
<th>No of Studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other Considerations</th>
<th>No of Participants</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (b,f)</td>
<td>randomized trials1</td>
<td>serious2</td>
<td>no serious inconsistency3</td>
<td>no serious indirectness4</td>
<td>serious5</td>
<td>none6</td>
<td>49/10,339 (0.4739%)</td>
<td>50/7,307 (0.6843%)</td>
<td>RR 0.6759 (0.4543 to 1.0057)</td>
<td>2,218 fewer (from 3,734 fewer to 39 more)</td>
</tr>
</tbody>
</table>

1 quasi-randomized
2 blinding and concealment were not clear
3 no heterogeneity exists; p-value for testing heterogeneity is 0.75 and $I^2=0\%$
4 the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
5 total sample size is large, but the total number of events is <300 (a threshold rule-of-thumb value)
6 insufficient number of studies to assess publication bias
## Summary of Findings Table for KQ1a – Breast Cancer Mortality (Aged 70–74)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Assumed Risk, Number per Million</th>
<th>Corresponding Risk, Number per Million</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Cancer Mortality for Ages 70-74</td>
<td>6,843</td>
<td>4,625</td>
<td>RR 0.6759 (0.4543 to 1.0057)</td>
<td>17,646 (2 studies(^1))</td>
<td>⊕⊕⊕⊕ low(^2,3,4,5,6)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>(3,109 to 6,882)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The assumed risk is the median control group risk. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

**CI:** Confidence Interval; **RR:** Risk Ratio

---

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

---

\(^1\) quasi-randomized
\(^2\) blinding and concealment were not clear
\(^3\) no heterogeneity exists; p-value for testing heterogeneity is 0.75 and I\(^2\)=0%
\(^4\) the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
\(^5\) total sample size is large, but the total number of events is <300 (a threshold rule-of-thumb value)
\(^6\) insufficient number of studies to assess publication bias
### Forest Plot for KQ1a – Breast Cancer Mortality (Aged 70–74)

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabár (E-County) 1995</td>
<td>23 5,033</td>
<td>31 4,868</td>
<td>54.6%</td>
<td>0.7176 [0.4190, 1.2289]</td>
<td></td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>26 5,306</td>
<td>19 2,439</td>
<td>45.4%</td>
<td>0.6290 [0.3488, 1.1343]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>10,339</strong></td>
<td><strong>7,307</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>0.6759 [0.4543, 1.0057]</strong></td>
<td><strong>0.6576 [0.4010, 1.0257]</strong></td>
</tr>
</tbody>
</table>

- **Total events**: 49 in Experimental and 50 in Control.
- **Heterogeneity**: \( \tau^2 = 0.00 \); \( \chi^2 = 0.10 \), df = 1 (\( P = 0.75 \)); \( I^2 = 0\% \)
- **Test for overall effect**: \( Z = 1.93 \) (\( P = 0.05 \))

\( \text{(M-H: Mantel-Haenszel)} \)
Evidence Set 3: KQ1a – Breast Cancer Mortality (Aged 50–69)

Does screening with mammography (film and digital) decrease breast cancer mortality for women aged 50–69?

- GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 50–69)
- Summary of Findings Table for KQ1a – Breast Cancer Mortality (Aged 50–69)
- Forest Plot for KQ1a – Breast Cancer Mortality (Aged 50–69): All 7 Studies Included
- Forest Plot for KQ1a – Breast Cancer Mortality (Aged 50–69): 2 Truly Randomized Studies Included
- Forest Plot for KQ1a – Breast Cancer Mortality (Aged 50–69): 5 Quasi-Randomized Studies Included
### GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 50–69)

#### Studies included:
- **a**: HIP; Habbema et al.66
- **b**: Kopparberg (W-County); Tabár et al.62
- **c2**: CNBSS-2; Miller et al.57
- **d**: Malmö; Nyström et al.71
- **e**: Stockholm; Nyström et al.71
- **f**: Östergötland (E-County); Nyström et al.71
- **g**: Göteborg; Bjurstam et al.63

<table>
<thead>
<tr>
<th>Quality Assessment</th>
<th>No of Participants</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Breast Cancer Mortality for Ages 50-69, All 7 Studies Included (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No of Studies</strong></td>
<td><strong>Design</strong></td>
<td><strong>Limitations</strong></td>
<td><strong>Inconsistency</strong></td>
<td><strong>Indirectness</strong></td>
</tr>
<tr>
<td>7 (a-g)</td>
<td>randomized trials</td>
<td>serious¹</td>
<td>no serious inconsistency²</td>
<td>no serious indirectness³</td>
</tr>
<tr>
<td><strong>Breast Cancer Mortality for Ages 50-69, 2 Truly Randomized Studies Included (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (c2,d)</td>
<td>randomized trials</td>
<td>no serious risk of bias</td>
<td>no serious inconsistency⁶</td>
<td>no serious indirectness³</td>
</tr>
<tr>
<td><strong>Breast Cancer Mortality for Ages 50-69, 5 Quasi-Randomized Studies Included (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 (a,b,e,f,g)</td>
<td>randomized trials</td>
<td>serious¹</td>
<td>no serious inconsistency⁷</td>
<td>no serious indirectness³</td>
</tr>
</tbody>
</table>

1. Blinding and concealment were not clear for five studies, so only two trials are considered truly randomized
2. No heterogeneity exists; p-value for testing heterogeneity is 0.12 and $I^2=41\%$
3. The question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
4. Total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
5. Insufficient number of studies to assess publication bias
6. No heterogeneity exists; p-value for testing heterogeneity is 0.25 and $I^2=24\%$
7. No heterogeneity exists; p-value for testing heterogeneity is 0.25 and $I^2=26\%$
## Summary of Findings Table for KQ1a – Breast Cancer Mortality (Aged 50–69)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative Comparative Risks* (95% CI)</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed Risk, Number per Million Control</td>
<td>Corresponding Risk, Number per Million Screening</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 50-69, All 7 Studies Included</td>
<td>6,449 (4,399 to 5,827)</td>
<td>RR 0.7850 (0.6821 to 0.9035)</td>
<td>250,274 (7 studies)</td>
<td>⊕⊕⊕⊝ moderate⁴,5</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>5,063 (4,399 to 5,827)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 50-69, 2 Truly Randomized Studies Included</td>
<td>7,309 (5,426 to 8,097)</td>
<td>RR 0.9069 (0.7424 to 1.1079)</td>
<td>73,047 (2 studies)</td>
<td>⊕⊕⊕⊕ high⁴,5,6</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>6,628 (5,426 to 8,097)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 50-69, 5 Quasi-Randomized Studies Included</td>
<td>6,050 (3,768 to 5,171)</td>
<td>RR 0.7296 (0.6228 to 0.8547)</td>
<td>177,227 (5 studies)</td>
<td>⊕⊕⊕⊕ moderate⁴,5,7</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>4,414 (3,768 to 5,171)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The **assumed risk** is the median control group risk. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio

---

### GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

---

1. blinding and concealment were not clear for five studies, so only two trials are considered truly randomized
2. no heterogeneity exists; p-value for testing heterogeneity is 0.12 and $I^2=41\%$
3. the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
4. total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
5. insufficient number of studies to assess publication bias
6. no heterogeneity exists; p-value for testing heterogeneity is 0.25 and $I^2=24\%$
7. no heterogeneity exists; p-value for testing heterogeneity is 0.25 and $I^2=26\%$
Forest Plot for KQ1a – Breast Cancer Mortality (Aged 50–69): All 7 Studies Included

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
</tr>
<tr>
<td>Bjurstam (Göteborg) 2003</td>
<td>54</td>
<td>9,926</td>
<td>103</td>
</tr>
<tr>
<td>Habbema (HIP) 1986</td>
<td>101</td>
<td>16,505</td>
<td>130</td>
</tr>
<tr>
<td>Miller (CNBSS-2) 2000</td>
<td>107</td>
<td>19,711</td>
<td>105</td>
</tr>
<tr>
<td>Nyström (Malmö) 2002</td>
<td>134</td>
<td>16,805</td>
<td>162</td>
</tr>
<tr>
<td>Nyström (Östergötland) 2002</td>
<td>117</td>
<td>23,584</td>
<td>137</td>
</tr>
<tr>
<td>Nyström (Stockholm) 2002</td>
<td>48</td>
<td>24,836</td>
<td>37</td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>78</td>
<td>23,701</td>
<td>69</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>135,068</td>
<td>115,206</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Total events: 639 (Experimental) 743 (Control)

Heterogeneity: Tau² = 0.01; Chi² = 10.11, df = 6 (P = 0.12); I² = 41%

Test for overall effect: Z = 3.38 (P = 0.0007)
(M-H: Mantel-Haenszel)
Forest Plot for KQ1a – Breast Cancer Mortality (Aged 50–69): 2 Truly Randomized Studies Included

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miller (CNBSS-2) 2000</td>
<td>Events 107</td>
<td>105</td>
<td>43.8%</td>
<td>1.0182 [0.7784, 1.3318]</td>
</tr>
<tr>
<td>Nyström (Malmö) 2002</td>
<td>Events 134</td>
<td>162</td>
<td>56.2%</td>
<td>0.8287 [0.6599, 1.0408]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>Events 241</td>
<td>267</td>
<td>100.0%</td>
<td>0.9069 [0.7424, 1.1079]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.01; Chi² = 1.31, df = 1 (P = 0.25); I² = 24%
Test for overall effect: Z = 0.96 (P = 0.34)
(M-H: Mantel-Haenszel)

Forest Plot for KQ1a – Breast Cancer Mortality (Aged 50–69): 5 Quasi-Randomized Studies Included

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bjurstam (Göteborg) 2003</td>
<td>Events 54</td>
<td>103</td>
<td>17.9%</td>
<td>0.8316 [0.5988, 1.1548]</td>
</tr>
<tr>
<td>Habbema (HIP) 1986</td>
<td>Events 101</td>
<td>130</td>
<td>25.2%</td>
<td>0.7769 [0.5996, 1.0067]</td>
</tr>
<tr>
<td>Nyström (Östergötland) 2002</td>
<td>Events 117</td>
<td>137</td>
<td>27.0%</td>
<td>0.8096 [0.6330, 1.0354]</td>
</tr>
<tr>
<td>Nyström (Stockholm) 2002</td>
<td>Events 48</td>
<td>37</td>
<td>11.6%</td>
<td>0.6768 [0.4410, 1.0386]</td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>Events 78</td>
<td>69</td>
<td>18.3%</td>
<td>0.5300 [0.3837, 0.7322]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>Events 398</td>
<td>476</td>
<td>100.0%</td>
<td>0.7296 [0.6228, 0.8547]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.01; Chi² = 5.39, df = 4 (P = 0.25); I² = 26%
Test for overall effect: Z = 3.90 (P < 0.0001)
(M-H: Mantel-Haenszel)
Evidence Set 4: KQ1a – Breast Cancer Mortality (Aged 50–59)

Does screening with mammography (film and digital) decrease breast cancer mortality for women aged 50–59?

- GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 50–59)
- Summary of Findings Table for KQ1a – Breast Cancer Mortality (Aged 50–59)
- Forest Plot for KQ1a – Breast Cancer Mortality (Aged 50–59)
### GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 50–59)

#### Studies included:
- **a**: HIP; Habbema et al.66
- **b**: Kopparberg (W-County); Tabár et al.62
- **c2**: CNBSS-2; Miller et al.57
- **d**: Malmö; Nyström et al.71
- **e**: Stockholm; Nyström et al.71
- **f**: Östergötland (E-County); Nyström et al.71
- **g**: Göteborg; Bjurstam et al.63

#### Quality Assessment

<table>
<thead>
<tr>
<th>No of Studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other Considerations</th>
<th>Screening</th>
<th>Control</th>
<th>Relative (95% CI)</th>
<th>Absolute per Million (Range)</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Cancer Mortality for Ages 50-59, All Studies (follow-up overall median 11.4 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 (a-g)</td>
<td>randomized trials¹</td>
<td>serious²</td>
<td>no serious inconsistency¹</td>
<td>no serious indirectness⁴</td>
<td>no serious imprecision⁵</td>
<td>none⁶</td>
<td>438/91,462 (0.4789%)</td>
<td>507/83,088 (0.6102%)</td>
<td>RR 0.8199 (0.6834 to 0.9835)</td>
<td>1,099 fewer (from 101 fewer to 1,932 fewer)</td>
<td>MODERATE</td>
<td>CRITICAL</td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 50-59, 2 Truly Randomized Trials (follow-up overall median 11.4 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (c2,d)</td>
<td>randomized trials⁷</td>
<td>no serious risk of bias</td>
<td>no serious inconsistency⁸</td>
<td>no serious indirectness⁴</td>
<td>no serious imprecision⁵</td>
<td>none⁶</td>
<td>195/28,996 (0.6725%)</td>
<td>195/29,016 (0.6720%)</td>
<td>RR 1.0013 (0.8216 to 1.2203)</td>
<td>9 more (from 1,199 fewer to 1,481 more)</td>
<td>HIGH</td>
<td>CRITICAL</td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 50-59, 5 Quasi-Randomized Studies (follow-up overall median 11.4 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 (a,b,e,f,g)</td>
<td>randomized trials</td>
<td>serious²</td>
<td>no serious inconsistency⁹</td>
<td>no serious indirectness⁴</td>
<td>no serious imprecision⁵</td>
<td>none⁶</td>
<td>243/62,466 (0.3890%)</td>
<td>312/54,072 (0.5770%)</td>
<td>RR 0.7336 (0.5859 to 0.9184)</td>
<td>1,537 fewer (from 471 fewer to 2,389 fewer)</td>
<td>MODERATE</td>
<td>CRITICAL</td>
</tr>
</tbody>
</table>

¹ five quasi-randomized and two truly randomized
² blinding and concealment were not clear for five studies, so only two trials are considered truly randomized
³ no heterogeneity exists; p-value for testing heterogeneity is 0.08 and $I^2=47\%$
⁴ the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
⁵ total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
⁶ insufficient number of studies to assess publication bias
⁷ truly randomized
⁸ no heterogeneity exists; p-value for testing heterogeneity is 0.86 and $I^2=0\%$
⁹ no heterogeneity exists; p-value for testing heterogeneity is 0.15 and $I^2=40\%$
## Summary of Findings Table for KQ1a – Breast Cancer Mortality (Aged 50–59)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative Comparative Risks* (95% CI)</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed Risk, Number per Million</td>
<td>Relative Effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Corresponding Risk, Number per Million</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Screening</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 50-59</td>
<td>6,102 (4,170 to 6,001)</td>
<td>RR 0.8199 (0.6834 to 0.9835)</td>
<td>174,550 (7 studies)</td>
<td>⊗蓥蓥蓥蓥 moderate^2,4,5,6</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 50-59, 2 Truly Randomized Trials</td>
<td>6,720 (5,522 to 8,201)</td>
<td>RR 1.0013 (0.8216 to 1.2203)</td>
<td>58,012 (2 studies)</td>
<td>⊗蓥蓥蓥蓥 high^3,5,6,8</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 50-59, 5 Quasi-Randomized Studies</td>
<td>5,770 (3,381 to 5,299)</td>
<td>RR 0.7336 (0.5859 to 0.9184)</td>
<td>116,538 (5 studies)</td>
<td>⊗蓥蓥蓥蓥 moderate^2,4,5,6,9</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The **assumed risk** is the median control group risk. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

**CI:** Confidence Interval; **RR:** Risk Ratio

**GRADE Working Group grades of evidence**

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

^1 five quasi-randomized and two truly randomized
^2 blinding and concealment were not clear for five studies, so only two trials are considered truly randomized
^3 no heterogeneity exists; p-value for testing heterogeneity is 0.08 and $I^2=47%$
^4 the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
^5 total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
^6 insufficient number of studies to assess publication bias
^7 truly randomized
^8 no heterogeneity exists; p-value for testing heterogeneity is 0.86 and $I^2=0%$
^9 no heterogeneity exists; p-value for testing heterogeneity is 0.15 and $I^2=40%$
## Forest Plot for KQ1a – Breast Cancer Mortality (Aged 50–59)

### Study

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
</tr>
<tr>
<td>Bjurstam (Göteborg) 2003</td>
<td>54</td>
<td>9,926</td>
<td>103</td>
</tr>
<tr>
<td>Habbema (HIP) 1986</td>
<td>77</td>
<td>12,855</td>
<td>97</td>
</tr>
<tr>
<td>Miller (CNBSS-2) 2000</td>
<td>107</td>
<td>19,711</td>
<td>105</td>
</tr>
<tr>
<td>Nyström (Malmö) 2002</td>
<td>88</td>
<td>9,285</td>
<td>90</td>
</tr>
<tr>
<td>Nyström (Östergötland) 2002</td>
<td>53</td>
<td>12,011</td>
<td>54</td>
</tr>
<tr>
<td>Nyström (Stockholm) 2002</td>
<td>25</td>
<td>15,946</td>
<td>24</td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>34</td>
<td>11,728</td>
<td>34</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td>91,462</td>
<td>83,088</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

- **Total events**: 438
- **Weight**: 15.6%
- **Chi²**: 11.31, df = 6 (P = 0.08); I² = 47%
- **Test for overall effect**: Z = 2.14 (P = 0.03)
- (M-H: Mantel-Haenszel)
Evidence Set 5: KQ1a – Breast Cancer Mortality (Aged 60–69)

Does screening with mammography (film and digital) decrease breast cancer mortality for women aged 60–69?

- GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 60–69)
- Summary of Findings Table for KQ1a – Breast Cancer Mortality (Aged 60–69)
- Forest Plot for KQ1a – Breast Cancer Mortality (Aged 60–69)
## GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (Aged 60–69)

**Studies included:**
- a: HIP; Habbema et al.\(^6\)
- b: Kopparberg (W-County); Tabár et al.\(^6\)
- d: Malmö; Nyström et al.\(^7\)
- e: Stockholm; Nyström et al.\(^7\)
- f: Östergötland (E-County); Nyström et al.\(^7\)

<table>
<thead>
<tr>
<th>Quality Assessment</th>
<th>Number of Participants</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Breast Cancer Mortality for Ages 60-69 (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 (a,b,d,e,f) randomized trials(^1)</td>
<td>201/43,606 (0.4609%) 236/32,118 (0.7348%)</td>
<td>RR 0.6850 (0.5665 to 0.8282) 2,315 fewer (from 1,262 fewer to 3,185 fewer)</td>
<td>MODERATE</td>
<td>CRITICAL</td>
</tr>
<tr>
<td><strong>Breast Cancer Mortality for Ages 60-69, 4 Quasi-Randomized Trials (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (a,b,c,e,f) randomized trials(^2)</td>
<td>155/36,086 (0.4295%) 164/24,603 (0.6666%)</td>
<td>RR 0.7026 (0.5630 to 0.8768) 1,982 fewer (from 821 fewer to 2,913 fewer)</td>
<td>MODERATE</td>
<td>CRITICAL</td>
</tr>
</tbody>
</table>

\(^1\) four quasi-randomized and one truly randomized  
\(^2\) blinding and concealment were not clear for five studies, so only two trials are considered truly randomized  
\(^3\) no heterogeneity exists; p-value for testing heterogeneity is 0.84 and \(I^2=0\%\)  
\(^4\) the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome  
\(^5\) the total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)  
\(^6\) insufficient number of studies to assess publication bias  
\(^7\) quasi-randomized  
\(^8\) no heterogeneity exists; p-value for testing heterogeneity is 0.74 and \(I^2=0\%\)
<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative Comparative Risks* (95% CI)</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed Risk, Number per Million Control</td>
<td>Corresponding Risk, Number per Million Screening</td>
<td>RR 0.6850 (0.5665 to 0.8282)</td>
<td>75,724 (5 studies1)</td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 60-69</td>
<td>7,348</td>
<td>5,033 (4,163 to 6,086)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6,666</td>
<td>4,683 (3,753 to 5,845)</td>
<td>RR 0.7026 (0.5630 to 0.8768)</td>
<td>60,689 (4 studies7)</td>
</tr>
<tr>
<td>Breast Cancer Mortality for Ages 60-69, 4 Quasi-Randomized Trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The assumed risk is the median control group risk. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence Interval; RR: Risk Ratio

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

1 four quasi-randomized and one truly randomized
2 blinding and concealment were not clear for five studies, so only two trials are considered truly randomized
3 no heterogeneity exists; p-value for testing heterogeneity is 0.84 and $I^2=0\%$
4 the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
5 the total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
6 insufficient number of studies to assess publication bias
7 quasi-randomized
8 no heterogeneity exists; p-value for testing heterogeneity is 0.74 and $I^2=0\%$
Forest Plot for KQ1a – Breast Cancer Mortality (Aged 60–69)

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
<td>Total</td>
</tr>
<tr>
<td>Habbema (HIP) 1986</td>
<td>24</td>
<td>3,650</td>
<td>33</td>
<td>3,650</td>
</tr>
<tr>
<td>Nyström (Malmö) 2002</td>
<td>46</td>
<td>7,520</td>
<td>72</td>
<td>7,515</td>
</tr>
<tr>
<td>Nyström (Östergötland) 2002</td>
<td>64</td>
<td>11,573</td>
<td>83</td>
<td>10,862</td>
</tr>
<tr>
<td>Nyström (Stockholm) 2002</td>
<td>23</td>
<td>8,890</td>
<td>13</td>
<td>4,536</td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>44</td>
<td>11,973</td>
<td>35</td>
<td>5,555</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td>43,606</td>
<td>32,118</td>
<td>100.0%</td>
<td>0.6850 [0.5665, 0.8282]</td>
</tr>
</tbody>
</table>

Test for overall effect: Z = 3.91 (P < 0.0001)
(M-H: Mantel-Haenszel)

Total events: 201 Experimental, 236 Control

Heterogeneity: Tau² = 0.00; Chi² = 1.44, df = 4 (P = 0.84); I² = 0%

0.1 0.2 0.5 1 2 5 10
Favours Experimental  Favours Control
Evidence Set 6: KQ1a – Breast Cancer Mortality (All Age Groups)

Does screening with mammography (film and digital) decrease breast cancer mortality for women in all age groups?

- GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (All Age Groups)
- Summary of Findings Table for KQ1a – Breast Cancer Mortality (All Age Groups)
- Forest Plot for KQ1a – Breast Cancer Mortality (All Age Groups)
GRADE Evidence Profile Table for KQ1a – Breast Cancer Mortality (All Age Groups)

Studies included:

a: HIP; Habbema et al.\textsuperscript{66}
b: Kopparberg (W-County); Tabár et al.\textsuperscript{62}
c1: CNBSS-1; Miller et al.\textsuperscript{56}
c2: CNBSS-2; Miller et al.\textsuperscript{57}
d: Malmö; Nyström et al.\textsuperscript{71}
e: Stockholm; Nyström et al.\textsuperscript{71}
f: Östergötland (E-County); Nyström et al.\textsuperscript{71}
g: Göteborg; Bjurstam et al.\textsuperscript{63}
h: AGE; Moss et al.\textsuperscript{61}

<table>
<thead>
<tr>
<th>Quality Assessment</th>
<th>No of Participants</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No of Studies</td>
<td>Design</td>
<td>Limitations</td>
<td>Inconsistency</td>
</tr>
<tr>
<td>Breast Cancer Mortality for All Age Groups (follow-up overall median 11.4 years)</td>
<td>9</td>
<td>randomized trials\textsuperscript{1}</td>
<td>serious\textsuperscript{2}</td>
<td>no serious inconsistency\textsuperscript{3}</td>
</tr>
</tbody>
</table>

\textsuperscript{1} five quasi-randomized and four truly randomized

\textsuperscript{2} blinding and concealment were not clear for five studies, so only four trials are considered truly randomized

\textsuperscript{3} no heterogeneity exists; p-value for testing heterogeneity is 0.13 and $I^2=36$

\textsuperscript{4} the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome

\textsuperscript{5} total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)

\textsuperscript{6} insufficient number of studies to assess publication bias
Summary of Findings Table for KQ1a – Breast Cancer Mortality (All Age Groups)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks* (95% CI)</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Cancer Mortality for All Age Groups</td>
<td>Assumed Risk, Number per Million Control</td>
<td>Corresponding Risk, Number per Million Screening</td>
<td>RR 0.8200 (0.7418 to 0.9065)</td>
<td>616,757 (9 studies¹) ⊕⊕⊕⊝ moderate²,³,⁴,⁵,⁶</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>4,433</td>
<td>3,635 (3,289 to 4,019)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The assumed risk is the median control group risk. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence Interval; RR: Risk Ratio

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

¹ five quasi-randomized and four truly randomized
² blinding and concealment were not clear for five studies, so only four trials are considered truly randomized
³ no heterogeneity exists; p-value for testing heterogeneity is 0.13 and $I^2=36$
⁴ the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
⁵ total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
⁶ insufficient number of studies to assess publication bias
### Forest Plot for KQ1a – Breast Cancer Mortality (All Age Groups)

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bjurstam (Göteborg) 2003</td>
<td>88</td>
<td>162</td>
<td>10.1%</td>
<td>0.7517 [0.5802, 0.9739]</td>
<td></td>
</tr>
<tr>
<td>Habbema (HIP) 1986</td>
<td>165</td>
<td>212</td>
<td>13.7%</td>
<td>0.7783 [0.6354, 0.9533]</td>
<td></td>
</tr>
<tr>
<td>Miller (CNBSS-2) 2000</td>
<td>107</td>
<td>105</td>
<td>9.6%</td>
<td>1.0182 [0.7784, 1.3318]</td>
<td></td>
</tr>
<tr>
<td>Miller (CNBSS-1) 2002</td>
<td>105</td>
<td>108</td>
<td>9.7%</td>
<td>0.9723 [0.7437, 1.2712]</td>
<td></td>
</tr>
<tr>
<td>Moss (AGE) 2006</td>
<td>105</td>
<td>251</td>
<td>12.0%</td>
<td>0.8303 [0.6614, 1.0425]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Malmö) 2002</td>
<td>190</td>
<td>231</td>
<td>14.6%</td>
<td>0.7887 [0.6514, 0.9549]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Östergötland) 2002</td>
<td>177</td>
<td>190</td>
<td>13.6%</td>
<td>0.9013 [0.7348, 1.1055]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Stockholm) 2002</td>
<td>82</td>
<td>50</td>
<td>6.5%</td>
<td>0.8790 [0.6186, 1.2490]</td>
<td></td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>126</td>
<td>104</td>
<td>10.1%</td>
<td>0.5834 [0.4502, 0.7559]</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong> (95% CI)</td>
<td><strong>298,043</strong></td>
<td><strong>318,714</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>0.8200 [0.7418, 0.9065]</strong></td>
<td></td>
</tr>
</tbody>
</table>

Total events: 1,413
Heterogeneity: Tau² = 0.01; Chi² = 12.51, df = 8 (P = 0.13); I² = 36%
Test for overall effect: Z = 3.88 (P = 0.0001)
(M-H: Mantel-Haenszel)
Evidence Set 7: KQ1a – All Cause Mortality

Does screening with mammography (film and digital) reduce all cause mortality?

- GRADE Evidence Profile Table for KQ1a – All Cause Mortality
- Summary of Findings Table for KQ1a – All Cause Mortality
- Forest Plot for KQ1a – All Cause Mortality (Aged 40–49)
- Forest Plot for KQ1a – All Cause Mortality (Aged 50–59)
## GRADE Evidence Profile Table for KQ1a – All Cause Mortality

**Studies included:**
- **c1:** CNBSS-1; Miller et al.\(^5\)\(^6\)
- **c2:** CNBSS-2; Miller et al.\(^5\)\(^7\)
- **h:** AGE trial; Moss et al.\(^6\)\(^1\)

<table>
<thead>
<tr>
<th>Quality Assessment</th>
<th>No of Participants</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No of Studies</td>
<td>Design</td>
<td>Limitations</td>
<td>Inconsistency</td>
</tr>
<tr>
<td><strong>All Cause Mortality for Ages 40-49 (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (c1,h)</td>
<td>randomized trials(^1)</td>
<td>no serious risk of bias</td>
<td>no serious inconsistency(^2)</td>
<td>no serious indirectness(^3)</td>
</tr>
<tr>
<td><strong>All Cause Mortality for Ages 50-59 (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (c2)</td>
<td>randomized trials(^1)</td>
<td>no serious risk of bias</td>
<td>no serious inconsistency(^6)</td>
<td>no serious indirectness(^3)</td>
</tr>
</tbody>
</table>

\(^1\) truly randomized  
\(^2\) no heterogeneity exists; p-value for testing heterogeneity is 0.65 and \(I^2=0\%\)  
\(^3\) the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome  
\(^4\) sample size is large and total number of events is > 300 (a threshold rule-of-thumb value)  
\(^5\) insufficient number of studies to assess publication bias  
\(^6\) single study; heterogeneity not applicable
## Summary of Findings Table for KQ1a – All Cause Mortality

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustative Comparative Risks* (95% CI)</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Cause Mortality for Ages 40-49</td>
<td>Assumed Risk, Number per Million Control</td>
<td>Corresponding Risk, Number per Million Screening</td>
<td>RR 0.9732 (0.9106 to 1.0402)</td>
<td>211,270 (2 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>18,067 (16,452 to 18,794)</td>
<td></td>
<td></td>
<td>⊕⊕⊕⊕ high</td>
</tr>
<tr>
<td>All Cause Mortality for Ages 50-59</td>
<td>Assumed Risk, Number per Million Control</td>
<td>Corresponding Risk, Number per Million Screening</td>
<td>RR 1.0629 (0.9598 to 1.1770)</td>
<td>39,405 (1 study)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td>35,036 (33,628 to 41,237)</td>
<td></td>
<td></td>
<td>⊕⊕⊕⊕ high</td>
</tr>
</tbody>
</table>

*The assumed risk is the median control group risk. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence Interval; RR: Risk Ratio

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

1 truly randomized
2 no heterogeneity exists; p-value for testing heterogeneity is 0.65 and I²=0%
3 the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
4 sample size is large and total number of events is > 300 (a threshold rule-of-thumb value)
5 insufficient number of studies to assess publication bias
6 single study; heterogeneity not applicable
Forest Plot for KQ1a – All Cause Mortality (Aged 40–49)

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental Events</th>
<th>Experimental Total</th>
<th>Control Events</th>
<th>Control Total</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miller (CNBSS-1) 2002</td>
<td>413</td>
<td>25,214</td>
<td>413</td>
<td>25,216</td>
<td>24.2%</td>
<td>1.0001 [0.8735, 1.1449]</td>
<td></td>
</tr>
<tr>
<td>Moss (AGE) 2006</td>
<td>960</td>
<td>53,884</td>
<td>1,975</td>
<td>106,956</td>
<td>75.8%</td>
<td>0.9648 [0.8938, 1.0414]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>79,098</td>
<td>132,172</td>
<td>1,373</td>
<td>2,388</td>
<td>100.0%</td>
<td>0.9732 [0.9106, 1.0402]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.00; Chi² = 0.20, df = 1 (P = 0.65); I² = 0%
Test for overall effect: Z = 0.80 (P = 0.42)
(M-H: Mantel-Haenszel)

Forest Plot for KQ1a – All Cause Mortality (Aged 50–59)

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental Events</th>
<th>Experimental Total</th>
<th>Control Events</th>
<th>Control Total</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miller (CNBSS-2) 2000</td>
<td>734</td>
<td>19,711</td>
<td>690</td>
<td>19,694</td>
<td>100.0%</td>
<td>1.0629 [0.9598, 1.1770]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>19,711</td>
<td>19,694</td>
<td>734</td>
<td>690</td>
<td>100.0%</td>
<td>1.0629 [0.9598, 1.1770]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Not applicable
Test for overall effect: Z = 1.17 (P = 0.24)
(M-H: Mantel-Haenszel)
Evidence Set 8: KQ1c – BSE and All Cause Mortality (Aged ≥40)

Does BSE practice decrease all cause mortality for women aged ≥40?

- GRADE Evidence Profile Table for KQ1c – BSE and All Cause Mortality (Aged ≥40)
- Summary of Findings Table for KQ1c – BSE and All Cause Mortality (Aged ≥40)
- Forest Plot for KQ1c – BSE and All Cause Mortality (Aged ≥40)
- Risk of Bias Table for KQ1c – BSE and All Cause Mortality (Aged ≥40)
GRADE Evidence Profile Table for KQ1c – BSE and All Cause Mortality (Aged ≥40)

Studies included:

i: Thomas\textsuperscript{73}  j: Semiglazov\textsuperscript{137}

<table>
<thead>
<tr>
<th>Study</th>
<th>No of Studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other Considerations</th>
<th>No of Participants</th>
<th>Effect (95% CI)</th>
<th>Absolute per Million (Range)</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i,j)</td>
<td>2</td>
<td>randomized trials</td>
<td>serious\textsuperscript{1}</td>
<td>no serious inconsistency\textsuperscript{2}</td>
<td>no serious indirectness\textsuperscript{3}</td>
<td>no serious imprecision\textsuperscript{4}</td>
<td>none\textsuperscript{5}</td>
<td>292/193,596 (0.1508%)</td>
<td>RR 0.9807 (0.8347 to 1.1524)</td>
<td>30 (from 254 fewer to 234 more)</td>
<td>MODERATE</td>
<td>CRITICAL</td>
</tr>
</tbody>
</table>

\textsuperscript{1} blinding and concealment were not clear
\textsuperscript{2} no heterogeneity exists; p-value for testing heterogeneity is 0.58 and $I^2=0$
\textsuperscript{3} the question addressed is the same for the evidence regarding the population, comparator and outcome
\textsuperscript{4} sample size is large and total number of events $\geq 300$ (a threshold rule-of-thumb) value
\textsuperscript{5} insufficient number of studies to assess publication bias
### Summary of Findings Table for KQ1c – BSE and All Cause Mortality (Aged ≥40)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative Comparative Risks* (95% CI)</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed Risk, Number per Million Control</td>
<td>Corresponding Risk, Number per Million Breast Self-Examination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality for Ages 40 Years and Older Follow-up: range 12 to 15 years</td>
<td>1,538</td>
<td>1,508 (1,284 to 1,772)</td>
<td>RR 0.9807 (0.8347 to 1.1524)</td>
<td>387,359 (2 studies)</td>
</tr>
</tbody>
</table>

*The **assumed risk** is the median control group risk. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence Interval; RR: Risk Ratio

**GRADE Working Group grades of evidence**

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are uncertain about the estimate.

1. blinding and concealment were not clear
2. no heterogeneity exists; p-value for testing heterogeneity is 0.58 and I²=0%
3. the question addressed is the same for the evidence regarding the population, comparator and outcome
4. sample size is large and total number of events >300 (a threshold rule-of-thumb) value
5. insufficient number of studies to assess publication bias
Forest Plot for KQ1c – BSE and All Cause Mortality (Aged ≥40)

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
<td>Total</td>
</tr>
<tr>
<td>Semiglazov 1992</td>
<td>157</td>
<td>60,617</td>
<td>167</td>
<td>60,678</td>
</tr>
<tr>
<td>Thomas 2002</td>
<td>135</td>
<td>132,979</td>
<td>131</td>
<td>133,085</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>193,596</td>
<td>193,763</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>

Total events 292

Heterogeneity: Tau² = 0.00; Chi² = 0.31, df = 1 (P = 0.58); I² = 0%

Test for overall effect: Z = 0.24 (P = 0.81)

(M-H: Mantel-Haenszel)
## Risk of Bias Table for KQ1c – BSE and All Cause Mortality (Aged ≥40)

<table>
<thead>
<tr>
<th>Author / Year</th>
<th>Randomization</th>
<th>Allocation Concealment</th>
<th>Blinding</th>
<th>Loss to Follow-up /Intention to Treat Analysis (ITT) Principle Observed or per Protocol Analysis</th>
<th>Other</th>
</tr>
</thead>
</table>
| **Semiglazov 1992**<sup>117</sup>  
Russia | Cluster randomized.  
Randomization units were 28 clinics in St. Petersburg, Russia.  
Randomization by computer (random digits at the WHO headquarters in Geneva). | Unclear – does the software conceal the allocation sequence? | Unclear (none reported) | Compliance went down over 4 years from 82% to 55.8%.  
Tracked breast cancer cases at the clinics and matched them to study participants. | |
| **Thomas 2002**<sup>73</sup>  
Shanghai | 519 factories were randomized to control (n=259) or intervention groups (260).  
~factory was the unit of randomization.  
~factories stratified by size and hospital affiliation.  
~method of randomization not stated. | Unclear | Those who determine deaths from breast cancer are blinded. | Follow-up done through factory medical clinic and home visits, and death/cancer registries.  
~ they estimate they missed 15 cases.  
~ the estimated number of missed cases was similar in both groups.  
~ 7.5 percent of women cut ties to their factory and were lost to the full 10 years of follow-up.  
~ the percentage is about the same in each group. (ITT) unclear. | One factory, mistakenly included as a control in preliminary article, was excluded in the final results.  
~2.6% of women in the instruction group were excluded after randomization  
~1.0% of women in the control group were excluded after randomization.  
~Author states that they didn’t think it would have an effect on results. |
Evidence Set 9: KQ2a – Harms of Screening/Mammography (Unnecessary Surgeries)

What are the harms associated with screening with mammography (unnecessary surgeries)?

- GRADE Evidence Profile Table for KQ2a – Harms of Screening/Mammography
- Summary of Findings Table for KQ2a – Harms of Screening/Mammography
GRADE Evidence Profile Table for KQ2a – Harms of Screening/Mammography (Unnecessary Surgeries)

**Source document:** Screening for breast cancer with mammography (Gøtzsche and Nielsen)\(^8\)

**Studies included:**

- **b:** Kopparberg (W-County); Tabár et al.\(^55\)
- **c1:** CNBSS-1; Miller et al.\(^56\)
- **c2:** CNBSS-2; Miller et al.\(^57\)
- **d:** Malmö; Nyström et al.\(^71\)
- **e:** Stockholm; Nyström et al.\(^71\)

<table>
<thead>
<tr>
<th>Quality Assessment</th>
<th>No of Participants</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Studies</td>
<td>Design</td>
<td>Limitations</td>
<td>Inconsistency</td>
<td>Indirectness</td>
</tr>
<tr>
<td><strong>Mastectomies and Lumpectomies, Includes 3 Truly Randomized Trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (c1,c2,d)</td>
<td>randomized trials</td>
<td>no serious risk of bias(^1)</td>
<td>no serious inconsistency(^2)</td>
<td>no serious indirectness(^3)</td>
</tr>
<tr>
<td><strong>Mastectomies and Lumpectomies, Includes 2 Quasi-randomized Trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (b,e)</td>
<td>randomized trials</td>
<td>serious(^6)</td>
<td>no serious inconsistency(^7)</td>
<td>no serious indirectness(^3)</td>
</tr>
<tr>
<td><strong>Mastectomies, Includes 3 Truly Randomized Trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (c1,c2,d)</td>
<td>randomized trials</td>
<td>no serious risk of bias(^1)</td>
<td>no serious inconsistency(^8)</td>
<td>no serious indirectness(^3)</td>
</tr>
<tr>
<td><strong>Mastectomies, Includes 2 Quasi-randomized Trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (b,e)</td>
<td>randomized trials</td>
<td>serious(^6)</td>
<td>no serious inconsistency(^9)</td>
<td>no serious indirectness(^3)</td>
</tr>
</tbody>
</table>

1. truly randomized  
2. no heterogeneity exists; p-value for testing heterogeneity is 0.87 and \(I^2=0\)%  
3. the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome  
4. total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)  
5. insufficient number of studies to assess publication bias  
6. blinding and concealment were not clear  
7. no heterogeneity exists; p-value for testing heterogeneity is 0.61 and \(I^2=0\)%  
8. no heterogeneity exists; p-value for testing heterogeneity is 0.65 and \(I^2=0\)%  
9. no heterogeneity exists; p-value for testing heterogeneity is 0.50 and \(I^2=0\)%
<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative Comparative Risks* (95% CI)</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mastectomies and Lumpectomies, Includes 3 Truly Randomized Trials</td>
<td>Assumed Risk, Number per Million Control: 16,371 (19,900 to 23,273)</td>
<td>RR 1.3146 (1.2156 to 1.4216)</td>
<td>132,321 (3 studies)</td>
<td>⊕⊕⊕⊕ high 1,2,3,4,5</td>
</tr>
<tr>
<td>Mastectomies and Lumpectomies, Includes 2 Quasi-randomized Trials</td>
<td>Assumed Risk, Number per Million Control: 8,662 (10,884 to 13,926)</td>
<td>RR 1.4213 (1.2565 to 1.6077)</td>
<td>118,158 (2 studies)</td>
<td>⊕⊕⊕⊕ moderate 3,4,5,6,7</td>
</tr>
<tr>
<td>Mastectomies, Includes 3 Truly Randomized Trials</td>
<td>Assumed Risk, Number per Million Control: 10,158 (10,977 to 13,455)</td>
<td>RR 1.1964 (1.0806 to 1.3246)</td>
<td>132,321 (3 studies)</td>
<td>⊕⊕⊕⊕ high 1,3,4,5,8</td>
</tr>
<tr>
<td>Mastectomies, Includes 2 Quasi-randomized Trials</td>
<td>Assumed Risk, Number per Million Control: 7,657 (8,096 to 10,585)</td>
<td>RR 1.2090 (1.0573 to 1.3824)</td>
<td>118,158 (2 studies)</td>
<td>⊕⊕⊕⊕ moderate 3,4,5,6,9</td>
</tr>
</tbody>
</table>

*The assumed risk is the median control group risk. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence Interval; RR: Risk Ratio

**GRADE Working Group grades of evidence**

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

1. truly randomized
2. no heterogeneity exists; p-value for testing heterogeneity is 0.87 and $I^2=0\%$
3. the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
4. total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
5. insufficient number of studies to assess publication bias
6. blinding and concealment were not clear
7. no heterogeneity exists; p-value for testing heterogeneity is 0.61 and $I^2=0\%$
8. no heterogeneity exists; p-value for testing heterogeneity is 0.65 and $I^2=0\%$
9. no heterogeneity exists; p-value for testing heterogeneity is 0.50 and $I^2=0\%$
Evidence Set 10: CQ4 – Optimal Mammography Screening Intervals

What are the optimal intervals for mammography screening?

- GRADE Evidence Profile Table for CQ4 – Optimal Mammography Screening Intervals
- Summary of Findings Table for CQ4 – Optimal Mammography Screening Intervals
- Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals < 24 Months for Ages 39-49
- Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 39-49
- Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 70-74
- Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals < 24 Months for Ages 50-69
- Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 50-69
- Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals < 24 Months for All Ages
- Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals ≥ 24 Months for All Ages
GRADE Evidence Profile Table for CQ4 – Optimal Mammography Screening Intervals

Included studies:
- a: HIP; Habbema et al.\textsuperscript{66}
- b: Kopparberg (W-County); Tabár et al.\textsuperscript{62}
- c1: CNBSS-1; Miller et al.\textsuperscript{56}
- c2: CNBSS-2; Miller et al.\textsuperscript{57}
- d: Malmö; Nyström et al.\textsuperscript{71}
- e: Stockholm; Nyström et al.\textsuperscript{71}
- f: Östergötland (E-County); Nyström et al.\textsuperscript{71}; for ages 70-74 year, Tabar et al.\textsuperscript{62}
- g: Göteborg; Bjurstam et al.\textsuperscript{63}
- h: AGE; Moss et al.\textsuperscript{61}

<table>
<thead>
<tr>
<th>No of Studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other Considerations</th>
<th>Screening</th>
<th>Control</th>
<th>Relative (95% CI)</th>
<th>Absolute per Million (Range)</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Breast Cancer Mortality for Screening Intervals &lt; 24 Months for Ages 39-49 (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 (a,c1,d,g,h)</td>
<td>randomized trials</td>
<td>no serious risk of bias\textsuperscript{1}</td>
<td>no serious inconsistency\textsuperscript{2}</td>
<td>no serious indirectness</td>
<td>no serious imprecision\textsuperscript{3}</td>
<td>none\textsuperscript{4}</td>
<td>361/118,130 (0.3056%)</td>
<td>566/172,408 (0.3283%)</td>
<td>RR 0.8247 (0.7215 to 0.9427)</td>
<td>575 fewer (from 188 fewer to 914 fewer)</td>
<td>☢☢☢☢</td>
<td>CRITICAL</td>
</tr>
<tr>
<td><strong>Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 39-49 (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (b,e,f)</td>
<td>randomized trials</td>
<td>serious\textsuperscript{5}</td>
<td>no serious inconsistency\textsuperscript{6}</td>
<td>no serious indirectness</td>
<td>serious\textsuperscript{8}</td>
<td>none\textsuperscript{4}</td>
<td>87/34,170 (0.2546%)</td>
<td>59/23,511 (0.2509%)</td>
<td>RR 1.0396 (0.7201 to 1.5008)</td>
<td>99 more (from 702 fewer to 1,257 more)</td>
<td>☢☢☢☢</td>
<td>LOW CRITICAL</td>
</tr>
<tr>
<td><strong>Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 70-74 (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (b,f)</td>
<td>randomized trials</td>
<td>serious\textsuperscript{5}</td>
<td>no serious inconsistency\textsuperscript{9}</td>
<td>no serious indirectness</td>
<td>serious\textsuperscript{10}</td>
<td>none\textsuperscript{4}</td>
<td>49/10,339 (0.4739%)</td>
<td>50/7,307 (0.6843%)</td>
<td>RR 0.6759 (0.4543 to 1.0057)</td>
<td>2,218 fewer (from 3,734 fewer to 39 more)</td>
<td>☢☢☢☢</td>
<td>LOW CRITICAL</td>
</tr>
<tr>
<td><strong>Breast Cancer Mortality for Screening Intervals &lt; 24 Months for Ages 50-69 (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (a,c2,d,g)</td>
<td>randomized trials</td>
<td>no serious risk of bias</td>
<td>no serious inconsistency\textsuperscript{11}</td>
<td>no serious indirectness</td>
<td>no serious imprecision\textsuperscript{1}</td>
<td>none\textsuperscript{4}</td>
<td>396/62,947 (0.6291%)</td>
<td>500/68,780 (0.7270%)</td>
<td>RR 0.8570 (0.7510 to 0.9781)</td>
<td>1,040 fewer (from 159 fewer to 1,810 fewer)</td>
<td>☢☢☢☢</td>
<td>CRITICAL</td>
</tr>
<tr>
<td><strong>Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 50-69 (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (b,e,f)</td>
<td>randomized trials</td>
<td>serious\textsuperscript{5}</td>
<td>no serious inconsistency\textsuperscript{12}</td>
<td>no serious indirectness</td>
<td>no serious imprecision\textsuperscript{2}</td>
<td>none\textsuperscript{4}</td>
<td>243/72,121 (0.3369%)</td>
<td>243/46,426 (0.5234%)</td>
<td>RR 0.6721 (0.5125 to 0.8815)</td>
<td>1,716 fewer (from 620 fewer to 2,552 fewer)</td>
<td>☢☢☢☢</td>
<td>MODERATE</td>
</tr>
<tr>
<td><strong>Breast Cancer Mortality for Screening Intervals &lt; 24 Months for All Ages (follow-up overall median 11.4 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 (a,c1,c2, d,g,h)</td>
<td>randomized trials</td>
<td>no serious risk of bias\textsuperscript{13}</td>
<td>no serious inconsistency\textsuperscript{14}</td>
<td>no serious indirectness</td>
<td>no serious imprecision\textsuperscript{3}</td>
<td>none\textsuperscript{4}</td>
<td>760/181,373 (0.4190%)</td>
<td>1,069/241,479 (0.4427%)</td>
<td>RR 0.8345 (0.7597 to 0.9165)</td>
<td>733 fewer (from 370 fewer to 1,064 fewer)</td>
<td>☢☢☢☢</td>
<td>CRITICAL</td>
</tr>
<tr>
<td>Breach Cancer Mortality for Screening Intervals ≥ 24 Months for All Ages (follow-up overall median 11.4 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>randomized trials</td>
<td>serious$^5$</td>
<td>serious$^{15}$</td>
<td>no serious indirectness$^7$</td>
<td>no serious imprecision$^7$</td>
<td>none$^4$</td>
<td>385/116,670 (0.3300%)</td>
<td>344/77,235 (0.4454%)</td>
<td>RR 0.7715 (0.5765 to 1.0326)</td>
<td>1.018 fewer (from 1.886 fewer to 145 more)</td>
<td>⊖ ⊖ ⊝ ⊝</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>two quasi-randomized and three truly randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>no heterogeneity exists; p-value for testing heterogeneity is 0.62 and $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>total sample size is large and the total number of events is &gt;300 (a threshold rule-of-thumb value)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>insufficient number of studies to assess publication bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>quasi-randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>no heterogeneity exists; p-value for testing heterogeneity is 0.31 and $I^2=15%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>total sample size is large but the total number of events is &lt;300 (a threshold rule-of-thumb value)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>no heterogeneity exists; p-value for testing heterogeneity is 0.75 and $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>the sample size is large but the total number of events is &lt;300 (a threshold rule-of-thumb value)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>no heterogeneity exists; p-value for testing heterogeneity is 0.52 and $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>no heterogeneity exists; p-value for testing heterogeneity is 0.12 and $I^2=52%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>four truly randomized and two quasi-randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>no heterogeneity exists; p-value for testing heterogeneity is 0.44 and $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>significant heterogeneity exists; p-value for testing heterogeneity is 0.03 and $I^2=72%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Summary of Findings Table for CQ4 – Optimal Mammography Screening Intervals

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative Comparative Risks* (95% CI)</th>
<th>Relative Effect (95% CI)</th>
<th>No of Participants (Studies)</th>
<th>Quality of the Evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed Risk, Number per Million Control</td>
<td>Corresponding Risk, Number per Million Screening</td>
<td>RR</td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Screening Intervals &lt; 24 Months for Ages 39-49</td>
<td>3,283</td>
<td>2,707 (2,369 to 3,095)</td>
<td>RR 0.8247 (0.7215 to 0.9427)</td>
<td>290,538 (5 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 39-49</td>
<td>2,509</td>
<td>2,609 (1,807 to 3,766)</td>
<td>RR 1.0396 (0.7201 to 1.5008)</td>
<td>57,681 (3 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 70-74</td>
<td>4,625 (3,109 to 6,882)</td>
<td>RR 0.6759 (0.4543 to 1.0057)</td>
<td>17,646 (2 studies)</td>
<td>⊕⊕ ⊝ ⊝ low3,5,7,9,10</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Screening Intervals &lt; 24 Months for Ages 50-69</td>
<td>7,270</td>
<td>6,230 (5,459 to 7,110)</td>
<td>RR 0.8570 (0.7510 to 0.9781)</td>
<td>131,727 (4 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 50-69</td>
<td>5,234</td>
<td>3,518 (2,682 to 4,614)</td>
<td>RR 0.6721 (0.5125 to 0.8815)</td>
<td>118,547 (3 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Screening Intervals &lt; 24 Months for All Ages</td>
<td>4,427</td>
<td>3,694 (3,363 to 4,057)</td>
<td>RR 0.8345 (0.7597 to 0.9165)</td>
<td>422,852 (6 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality for Screening Intervals ≥ 24 Months for All Ages</td>
<td>4,454</td>
<td>3,436 (2,568 to 4,599)</td>
<td>RR 0.7715 (0.5765 to 1.0326)</td>
<td>193,905 (3 studies)</td>
</tr>
<tr>
<td>Follow-up: overall median 11.4 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The assumed risk is the median control group risk. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

**CI:** Confidence Interval; **RR:** Risk Ratio

**GRADE Working Group grades of evidence**

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

1. two quasi-randomized and three truly randomized
2. no heterogeneity exists; p-value for testing heterogeneity is 0.62 and I²=0%
3. total sample size is large and the total number of events is >300 (a threshold rule-of-thumb value)
4. insufficient number of studies to assess publication bias
5. quasi-randomized
no heterogeneity exists; p-value for testing heterogeneity is 0.31 and $I^2=15$
the question addressed is the same for the evidence regarding the population, intervention, comparator and outcome
total sample size is large but the total number of events is <300 (a threshold rule-of-thumb value)
no heterogeneity exists; p-value for testing heterogeneity is 0.75 and $I^2=0$
the sample size is large but the total number of events is <300 (a threshold rule-of-thumb value)
no heterogeneity exists; p-value for testing heterogeneity is 0.52 and $I^2=0$
four truly randomized and two quasi-randomized
no heterogeneity exists; p-value for testing heterogeneity is 0.44 and $I^2=0$
significant heterogeneity exists; p-value for testing heterogeneity is 0.03 and $I^2=72$
### Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals < 24 Months for Ages 39–49

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental Events</th>
<th>Total</th>
<th>Control Events</th>
<th>Total</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bjurstam (Göteborg) 2003</td>
<td>34</td>
<td>11,724</td>
<td>59</td>
<td>14,217</td>
<td>10.1%</td>
<td>0.6988 [0.4586, 1.0649]</td>
<td></td>
</tr>
<tr>
<td>Habbema (HIP) 1986</td>
<td>64</td>
<td>13,740</td>
<td>82</td>
<td>13,740</td>
<td>16.8%</td>
<td>0.7805 [0.5633, 1.0814]</td>
<td></td>
</tr>
<tr>
<td>Miller (CNBSS-1) 2002</td>
<td>105</td>
<td>25,214</td>
<td>108</td>
<td>25,216</td>
<td>24.9%</td>
<td>0.9723 [0.7437, 1.2712]</td>
<td></td>
</tr>
<tr>
<td>Moss (AGE) 2006</td>
<td>105</td>
<td>53,884</td>
<td>251</td>
<td>106,956</td>
<td>34.5%</td>
<td>0.8303 [0.6614, 1.0425]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Malmö) 2002</td>
<td>53</td>
<td>13,568</td>
<td>66</td>
<td>12,279</td>
<td>13.7%</td>
<td>0.7267 [0.5067, 1.0424]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>118,130</strong></td>
<td><strong>118,130</strong></td>
<td><strong>172,408</strong></td>
<td><strong>172,408</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>0.8247 [0.7215, 0.9427]</strong></td>
<td></td>
</tr>
</tbody>
</table>

Total events: 361

Heterogeneity: Tau² = 0.00; Chi² = 2.63, df = 4 (P = 0.62); I² = 0%

Test for overall effect: Z = 2.83 (P = 0.005)

(M-H: Mantel-Haenszel)

---

### Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 39–49

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental Events</th>
<th>Total</th>
<th>Control Events</th>
<th>Total</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nyström (Östergötland) 2002</td>
<td>31</td>
<td>10,285</td>
<td>30</td>
<td>10,459</td>
<td>43.0%</td>
<td>1.0508 [0.6366, 1.7346]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Stockholm) 2002</td>
<td>34</td>
<td>14,303</td>
<td>13</td>
<td>8,021</td>
<td>28.7%</td>
<td>1.4667 [0.7745, 2.7775]</td>
<td></td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>22</td>
<td>9,582</td>
<td>16</td>
<td>5,031</td>
<td>28.3%</td>
<td>0.7219 [0.3795, 1.3734]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>34,170</strong></td>
<td><strong>23,511</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>23,511</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>1.0396 [0.7201, 1.5008]</strong></td>
<td></td>
</tr>
</tbody>
</table>

Total events: 87

Heterogeneity: Tau² = 0.02; Chi² = 2.36, df = 2 (P = 0.31); I² = 15%

Test for overall effect: Z = 0.21 (P = 0.84)

(M-H: Mantel-Haenszel)
Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 70-74

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental Events</th>
<th>Experimental Total</th>
<th>Control Events</th>
<th>Control Total</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabár (E-County) 1995</td>
<td>23</td>
<td>5,033</td>
<td>31</td>
<td>4,868</td>
<td>54.6%</td>
<td>0.7176 [0.4190, 1.2289]</td>
<td></td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>26</td>
<td>5,306</td>
<td>19</td>
<td>2,439</td>
<td>45.4%</td>
<td>0.6290 [0.3488, 1.1343]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>10,339</td>
<td>7,307</td>
<td>100.0%</td>
<td></td>
<td></td>
<td>0.6759 [0.4543, 1.0057]</td>
<td></td>
</tr>
<tr>
<td>Total events</td>
<td>49</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.00; Chi² = 0.10, df = 1 (P = 0.75); I² = 0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 1.93 (P = 0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M-H: Mantel-Haenszel)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals < 24 Months for Ages 50–69

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental Events</th>
<th>Experimental Total</th>
<th>Control Events</th>
<th>Control Total</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bjurstam (Göteborg) 2003</td>
<td>54</td>
<td>9,926</td>
<td>103</td>
<td>15,744</td>
<td>16.2%</td>
<td>0.8316 [0.5988, 1.1548]</td>
<td></td>
</tr>
<tr>
<td>Habbema (HIP) 1986</td>
<td>101</td>
<td>16,505</td>
<td>130</td>
<td>16,505</td>
<td>26.0%</td>
<td>0.7769 [0.5996, 1.0067]</td>
<td></td>
</tr>
<tr>
<td>Miller (CNBSS-2) 2000</td>
<td>107</td>
<td>19,711</td>
<td>105</td>
<td>19,694</td>
<td>24.2%</td>
<td>1.0182 [0.7784, 1.3318]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Malmö) 2002</td>
<td>134</td>
<td>16,805</td>
<td>162</td>
<td>16,837</td>
<td>33.6%</td>
<td>0.8287 [0.6599, 1.0408]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>62,947</td>
<td>68,780</td>
<td>100.0%</td>
<td></td>
<td></td>
<td>0.8570 [0.7510, 0.9781]</td>
<td></td>
</tr>
<tr>
<td>Total events</td>
<td>396</td>
<td>500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.00; Chi² = 2.25, df = 3 (P = 0.52); I² = 0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.29 (P = 0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M-H: Mantel-Haenszel)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals ≥ 24 Months for Ages 50–69

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
<td>Total</td>
</tr>
<tr>
<td>Nyström (Östergötland) 2002</td>
<td>117</td>
<td>23,584</td>
<td>137</td>
<td>22,357</td>
</tr>
<tr>
<td>Nyström (Stockholm) 2002</td>
<td>48</td>
<td>24,836</td>
<td>37</td>
<td>12,957</td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>78</td>
<td>23,701</td>
<td>69</td>
<td>11,112</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>72,121</strong></td>
<td><strong>46,426</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>0.6721 [0.5125, 0.8815]</strong></td>
</tr>
<tr>
<td>Total events</td>
<td>243</td>
<td>243</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.03; Chi² = 4.19, df = 2 (P = 0.12); I² = 52%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.87 (P = 0.004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M-H: Mantel-Haenszel)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals < 24 Months for All Ages

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
<td>Total</td>
</tr>
<tr>
<td>Bjurstam (Göteborg) 2003</td>
<td>88</td>
<td>21,650</td>
<td>162</td>
<td>29,961</td>
</tr>
<tr>
<td>Habbema (HIP) 1986</td>
<td>165</td>
<td>30,245</td>
<td>212</td>
<td>30,245</td>
</tr>
<tr>
<td>Miller (CNBSS-2) 2000</td>
<td>107</td>
<td>19,711</td>
<td>105</td>
<td>19,694</td>
</tr>
<tr>
<td>Miller (CNBSS-1) 2002</td>
<td>105</td>
<td>25,214</td>
<td>108</td>
<td>25,216</td>
</tr>
<tr>
<td>Moss (AGE) 2006</td>
<td>105</td>
<td>53,884</td>
<td>251</td>
<td>106,956</td>
</tr>
<tr>
<td>Nyström (Malmö) 2002</td>
<td>190</td>
<td>30,669</td>
<td>231</td>
<td>29,407</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>181,373</strong></td>
<td><strong>241,479</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>0.8345 [0.7597, 0.9165]</strong></td>
</tr>
<tr>
<td>Total events</td>
<td>760</td>
<td>1,069</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.00; Chi² = 4.77, df = 5 (P = 0.44); I² = 0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 3.78 (P = 0.0002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M-H: Mantel-Haenszel)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Forest Plot for CQ4 – Breast Cancer Mortality for Screening Intervals ≥ 24 Months for All Ages

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental Events</th>
<th>Total</th>
<th>Control Events</th>
<th>Total</th>
<th>Weight</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nyström (Östergötland) 2002</td>
<td>177</td>
<td>38,942</td>
<td>190</td>
<td>37,675</td>
<td>38.0%</td>
<td>0.9013 [0.7348, 1.1055]</td>
<td></td>
</tr>
<tr>
<td>Nyström (Stockholm) 2002</td>
<td>82</td>
<td>39,139</td>
<td>50</td>
<td>20,978</td>
<td>27.8%</td>
<td>0.8790 [0.6186, 1.2490]</td>
<td></td>
</tr>
<tr>
<td>Tabár (W-County) 1995</td>
<td>126</td>
<td>38,589</td>
<td>104</td>
<td>18,582</td>
<td>34.1%</td>
<td>0.5834 [0.4502, 0.7559]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>116,670</strong></td>
<td><strong>38,942</strong></td>
<td><strong>77,235</strong></td>
<td><strong>100.0%</strong></td>
<td></td>
<td><strong>0.7715 [0.5765, 1.0326]</strong></td>
<td></td>
</tr>
</tbody>
</table>

Total events: 385 344

Heterogeneity: Tau² = 0.05; Chi² = 7.20, df = 2 (P = 0.03); I² = 72%

Test for overall effect: Z = 1.74 (P = 0.08)
(M-H: Mantel-Haenszel)
Appendix 8: List of Studies Excluded at Full Text Screening

Not average-risk population


**Study is not about mammography, CBE, or BSE**


**Outcome not mortality, harm, or cost**


Excluded by study design


Appendix 9: List of External Reviewers – Protocol

E. Breslau  National Cancer Institute, Bethesda, USA
A. Chiarelli  Cancer Care Ontario, Toronto, Canada
J. Sussman  Juravinski Cancer Centre, Hamilton, Canada
Anonymous

We wish to acknowledge and thank these individuals for their input in the review protocol.
Appendix 10: List of External Reviewers – Evidence Synthesis

E. Breslau  National Cancer Institute, Bethesda, USA
H. Bryant  Canadian Partners Against Cancer, Toronto, Canada
A. Chiarelli  Cancer Care Ontario, Toronto, Canada
K. Johnson  Public Health Agency of Canada, Ottawa, Canada
G. Kruger  Radiology Consultants Associated, Calgary, Canada
V. Mai  Cancer Care Ontario, Toronto, Canada

We wish to acknowledge and thank these individuals for their input in this evidence review.
Acknowledgements

We would like to thank the following staff members for their work on this review:

Maureen Rice – Librarian

Sharon Peck-Reid – Research Assistant

We gratefully acknowledge the support of Canadian Institute for Health Research for funds to support the McMaster Evidence Review and Synthesis Centre for this systematic review.